

REMARKS

Applicants request entry of the amendments and reconsideration of the application.

Amendment to the Specification

The amendment to the specification corrects typographical or inadvertent errors and does not add new matter. As noted in the Appendix A, which shows the changes made, the errors relate to the abbreviated titles of journal references and the form and completeness of the citations to these journal references.

Claims 17-36 have been canceled and new claims 37-76 have been added.

The amendments to the claims reflect the selection of SEQ ID NO: 17 and 18 for examination, as required by the Examiner in his objections to the claims (see page 4 of Paper No. 19). Applicants specifically reserve the right to seek patents for all the sequences and subject matter disclosed in the application and original claims. The changes in the recitation of the polynucleotides of the claims are made to better reflect the disclosures in the specification and in the priority document. The amendments are not made for reasons of patentability.

As shown below, no new matter enters by the new claims and the specification enables the new claims.

The new claims are directed to transcription factors that have at least 42% amino acid sequence identity with the AP2 transcription factor of SEQ ID NO: 18, and have enhanced plant disease tolerance due to changed expression or activity of the transcription factor. These claims are disclosed and enabled by the text of the specification at, for example, page 17, lines 6-9; page 13, lines 13-23; page 22, lines 34-35; page 23, lines 36-37 and continued on page 24, lines 1-4; page 25, lines 6-11; at page 11, lines 21-22, which discloses "[t]ranscription factors that are homologs of the disclosed sequences will typically share at least 40% amino acid sequence identity"; and at page 25, lines 4-9, which discloses that G19 (SEQ ID NO:22) expression was significantly induced upon infection by the fungal pathogen *Erysiphe* and G19-overexpressing plants were more tolerant to *Erysiphe* infection; and by Figure 1, which discloses the AP2-family transcription factors G1006 (SEQ ID NO: 34), G19 (SEQ ID NO:22), G4 (SEQ ID NO:12) and

G25 (SEQ ID NO: 34), which, as shown below, share 60%, 42%, 42%, and 54% amino acid sequence identity with G28 (SEQ ID NO:18), respectively.

As shown in the attached Appendix B providing experimental expression data, G1006 and G25 expression is induced in response to exposure to different disease agents. As shown in the executed declaration of Dr. Jacqueline Heard (the Heard Declaration, attached as Appendix C), G1006, G19, G4 and G25 expression has been shown to be up-regulated by *Fusarium*, *Erysiphe*, *Botrytis* exposure or infection, and methyl jasmonate in microarray analyses (see in particular the Heard Declaration, Exhibits A through G of the Declaration). As disclosed in the specification and stated in the Heard Declaration, the jasmonic acid signal transduction pathway is involved in the regulation of the defense response; the regulation of G4 expression by both hormones, its induction upon *Erysiphe orontii* infection, as well as the preliminary data indicating that increased tolerance to that pathogen is conferred by G19 overexpression, suggest that these transcription factors control the defense response. Thus, evidence is provided that shows AP2 family transcription factors with at least 42% identity to SEQ ID NO:18 (G28) either induce enhanced tolerance to plant disease or provide enhanced tolerance based on sequence similarity and expression profiles, thus likely endowing a plant with increased pathogen resistance (as would be recognized by one skilled in the art; see the Heard declaration, sections 8 to 9).

As also shown in Heard Declaration, the function of SEQ ID NOs:17 and 18 (G28) was more recently analyzed using transgenic plants in which this gene was expressed under the control of the 35S promoter. G28 overexpressing lines were shown to be more tolerant to infections by *Erysiphe orontii*, *Sclerotinia sclerotiorum* or *Botrytis cinerea*. Thus, the disclosure in the specification, that the transcription factors recited in the claims, including G28, could be used to produce plants that are more tolerant to disease, has been confirmed by experimentation. (See, for example, page 7, lines 22-28, page 17, lines 10-17, page 19, line 31 through page 21, line 28, page 24, line 9 through page 25, line 13.)

The new claims are also directed to transcription factors that comprise a conserved domain of a plant AP2 transcription factor wherein the conserved domain comprises one of an amino acid sequence of residues 145-213 of the AP2 transcription factor of SEQ ID NO: 18, or

an amino acid sequence having at least 84% identity to residues 145-213 of the AP2 transcription factor of SEQ ID NO: 18, or an amino acid sequence of residues 145-213 of the AP2 transcription factor of SEQ ID NO: 18 having one or more conservative substitutions, deletions, or insertions. Applicants clearly identified amino acid residues 145-213 of the AP2 transcription factor of SEQ ID NO: 18 in Figure 1a (see SEQ ID NO: 18, G28, in particular). Also, the specification at page 10, lines 13-14, specifically notes the conservative substitutions, deletions, and insertions of which one of skill in the art would be aware. Further support for residues that may be substituted and with which residues they may be substituted is disclosed in Table 1 at page 46 of provisional Application serial No. 60/125,814, from which the instant application claims priority. In addition, fragments of domains of transcription factor sequences are specifically disclosed at page 4, lines 32-37, and in Figure 1.

Applicants also note that SEQ ID NO: 22 (G19) is disclosed in the specification as being significantly induced upon infection by the fungal pathogen *Erysiphe orontii* (see specification at page 25, lines 4-9). As shown in the accompanying BLAST report (Appendix F), the conserved domains of G28 (SEQ ID NO: 18) show substantial sequence identity to G19 (SEQ ID NO: 22), namely 60% identity (see Appendix F: BLAST report, page number listed as 12 or 28; handwritten notes regarding % identity provided by the applicants). Applicants have inserted this identity language into the new claims. The use of identity ranges is well known in the art and is specifically supported by the specification at, for example, page 11, lines 21-23. Thus, one of skill in the art would recognize the ability of the G19 sequence (SEQ ID NO: 22) and the closely related G28 sequence (SEQ ID NO: 18) to produce transgenic plants with at least pathogen stress resistance or disease resistance.

The BLAST report also shows that conserved domains of G4 (SEQ ID NO: 12) with 62% identity to SEQ ID NO: 18, and G1006 (SEQ ID NO: 34) with 97% identity to SEQ ID NO: 18, and G25 (SEQ ID NO: 82) with 69% identity, show substantial sequence identity to the conserved domain of G28 (SEQ ID NO: 18) (see handwritten notes regarding % identity provided by the applicants, particularly at pages 4, 10, and 13, respectively). The BLAST analysis method is fully disclosed in the specification at page 11, lines 3-20; at page 19, lines 2-8; and at page 25, lines 32-37. In addition, the specification discloses the function and utility of

these polynucleotide and polypeptide sequences (see specification at pages 1 through 3, and pages 5 through 8, where the different polynucleotide and polypeptide sequences and their use are disclosed). Accordingly, applicants' specification, in light of this evidence, clearly demonstrates the range of polynucleotide sequences that could be used in the new methods and transgenic plants.

To further demonstrate that the functional characteristics recited in the new claims would have been understood to be present with the recited polynucleotides, and/or easily identified in transgenic plants comprising these polynucleotides, applicants also enclose a copy of Table 1 of U.S. provisional application serial no. 60/125,814, at page 46 (see Appendix D, p 45, lines 15-32; page 46, Table 1). The present application claims priority to this provisional application. Table 1 shows amino acid residues of a polypeptide sequence that may be substituted with other amino acids and are restricted to conservative substitutions.

The conserved domains of AP2-family transcription factors are more nearly identical in their sequences, with G1006 (SEQ ID NO: 34), G19 (SEQ ID NO: 22), G4 (SEQ ID NO: 12) and G25 (SEQ ID NO: 82) sharing 98%, 72%, 69%, and 72% sequence identity, respectively, with the conserved domain of G28 (please see Appendix E, showing each of these comparisons and the sequence identity). Thus, structurally and functionally related AP2 family transcription factors share a minimum of 69% sequence identity in their conserved domains, which supports new claims directed to conserved domains.

Support for claim 37 is provided by claim 1 as filed (transgenic plant comprising recombinant polynucleotide altering disease resistance or tolerance), and the specification at page 6, lines 29-31 and page 17, lines 10-23 (expression of instant AP2 transcription factors leads to enhanced tolerance to microbial, fungal, nematode or viral diseases, pathogens and pests), and by page 11, lines 21-29 (transcription factors that are homologs may share at least 40% sequence identity).

Support for claim 38 may be found in Figure 1a (conserved domain comprises residues 145-213), and the specification at page 11, lines 21-29 (transcription factors that are homologs may share at least 40% sequence identity), and on page 6, lines 4-6 ("conserved domain" refers

to a polypeptide fragment that is more conserved at a sequence level than other fragments). The specification at page 10, lines 12-26 and page 10, lines 13-14 also notes conservative substitutions, deletions, and insertions of which one of skill in the art would be aware. Further support for residues that may be substituted and with which residues they may be substituted is disclosed in Table 1 at page 46 of provisional Application serial No. 60/125,814, from which the instant application claims priority.

Support for new claim 39 is provided by the specification at, for example, page 5, line 9, and page 16, line 9 through page 17, line 24.

Support for new claim 40 is provided by claim 3 as filed and on, for example, page 3, lines 24-26, on page 13, lines 13-19, and page 14, lines 13-18.

Support for new claim 41 is provided by claim 3 as filed and on page 13, lines 15-19.

Support for new claim 42 is provided by claim 4 as filed and on page 13, lines 16-17 and page 13, line 20 through page 14, line 12.

Support for new claim 43 is provided on, for example, page 17, lines 10-11 (utility for increasing tolerance or resistance to pathogens and pests), page 6, lines 29-31 and page 17, lines 10-23 (expression of instant AP2 transcription factors leads to enhanced tolerance to, fungal, and other pathogens), on page 19, line 32 through page 21, line 28 (those genes induced after exposure to biotrophic fungal pathogens), and on page 24, line 9 to page 25, line 13 (analysis of *Arabidopsis* T2 progeny plants for pathogen resistance or pathogen tolerance).

Support for new claim 44 is provided on page 7, lines 22-28, page 17, lines 10-17, page 19, line 31 through page 21, line 28, page 24, line 9 through page 25, line 13. Statements in the specification indicating that G28 can be used to confer tolerance to fungal pathogens have been confirmed by experimentation (see description in the Heard Declaration, above and below).

Support for new claim 45 is provided by claim 5 as filed (method for altering the disease tolerance or resistance of a plant), and, for example, on page 7, line 35 (modulating a plant's response to disease), page 17, lines 10-23 (increasing tolerance or resistance to pathogens and pests), and on page 11, lines 21-29 (transcription factors that are homologs may share at least 40% sequence identity).

Support for new claim 46 is provided in Figure 1a (conserved domain comprises residues 145-213), on page 11, lines 21-29 (transcription factors that are homologs may share at least 40% sequence identity), on page 6, lines 4-6 ("conserved domain" refers to a polypeptide fragment that is more conserved at a sequence level than other fragments), page 6, lines 29-31 and page 17, lines 10-23 (expression of instant AP2 transcription factors leads to enhanced tolerance to microbial, fungal, nematode or viral diseases, pathogens and pests). The specification at page 10, lines 12-26 and page 10, lines 13-14 also notes conservative substitutions, deletions, and insertions of which one of skill in the art would be aware. Further support for residues that may be substituted and with which residues they may be substituted is disclosed in Table 1 at page 46 of provisional Application serial No. 60/125,814, from which the instant application claims priority.

Support for new claim 47 is provided by, for example, page 5, line 9, page 6, lines 28-31, and page 17, lines 10-23.

Support for new claim 48 is provided by claim 8 as filed and on, for example, page 3, lines 24-26, on page 13, lines 13-19, and page 14, lines 13-18.

Support for new claim 49 is provided by claim 8 as filed and on page 13, lines 15-19.

Support for new claim 50 is provided by claim 9 as filed and on page 13, lines 16-17 and page 13, line 20 through page 14, line 12.

Support for new claim 51 is provided on, for example, page 17, lines 10-11 (increasing tolerance or resistance to pathogens and pests), page 6, lines 29-31 and page 17, lines 10-23 (expression of instant AP2 transcription factors leads to enhanced tolerance to, fungal, and other pathogens), on page 19, lines 32 through page 21, line 28 (those genes induced after exposure to biotrophic fungal pathogens), and on page 24, line 9 to page 25, line 13 (analysis of *Arabidopsis* T2 progeny plants for pathogen resistance or pathogen tolerance).

Support for new claims 52 is provided on page 7, lines 22-28, page 17, lines 10-17, page 19, line 31 through page 21, line 28, page 24, line 9 through page 25, line 13. Statements in the specification indicating that G28 can be used to confer tolerance to fungal pathogens have been confirmed by experimentation (see description in the Heard Declaration, above and below).

Support for new claim 53 is provided by claim 10 as filed (method for altering the expression levels of at least one gene in a plant), and for example, on page 2, lines 19-23 and page 8, lines 29-34 (producing transgenic plants with modified expression levels of at least one gene), on page 17, lines 10-23 (increasing tolerance or resistance to pathogens and pests), page 6, lines 29-31 and page 17, lines 10-23 (expression of instant AP2 transcription factors leads to enhanced tolerance to microbial, fungal, nematode or viral diseases, pathogens and pests), and on page 11, lines 21-29 (transcription factors that are homologs may share at least 40% sequence identity).

Support for new claim 54 may be found in Figure 1a (conserved domain comprises residues 145-213), on page 11, lines 21-29 (transcription factors that are homologs may share at least 40% sequence identity), and on page 6, lines 4-6 (“conserved domain” refers to a polypeptide fragment that is more conserved at a sequence level than other fragments). The specification at page 10, lines 12-26 and page 10, lines 13-14 also notes conservative substitutions, deletions, and insertions of which one of skill in the art would be aware. Further support for which residues may be substituted and with which residues they may be substituted is disclosed in Table 1 at page 46 of provisional Application serial No. 60/125,814, from which the instant application claims priority.

Support for new claim 55 is provided by, for example, page 5, line 9, page 6, lines 28-31, and page 12, line 27 through page 17, line 23.

Support for new claim 56 is provided by claim 12 as filed and on, for example, page 3, lines 24-26, on page 13, lines 13-19, and page 14, lines 13-18.

Support for new claim 57 is provided by claim 12 as filed and on page 13, lines 15-19.

Support for new claim 58 is provided by claim 13 as filed and on page 13, lines 16-17 and page 13, line 20 through page 14, line 12.

Support for new claim 59 is provided on, for example, page 17, lines 10-11 (utility for increasing tolerance or resistance to pathogens and pests), page 6, lines 29-31 and page 17, lines 10-23 (expression of instant AP2 transcription factors leads to enhanced tolerance to, fungal, and other pathogens), on page 19, lines 32 through page 21, line 28 (those genes induced after

exposure to biotrophic fungal pathogens), and on page 24, line 9 to page 25, line 13 (analysis of *Arabidopsis* T2 progeny plants for pathogen resistance or pathogen tolerance).

Support for new claim 60 is provided on page 7, lines 22-28, page 17, lines 10-17, page 19, line 31 through page 21, line 28, page 24, line 9 through page 25, line 13. Statements in the specification indicating that G28 can be used to confer tolerance to fungal pathogens have been confirmed by experimentation (see description in the Heard Declaration, above and below).

Support for new claim 61 can be found in claim 1 as filed and page 6, lines 29-31 (transgenic plant comprising recombinant polynucleotide altering disease resistance or tolerance), page 17, lines 10-23 (expression of instant AP2 transcription factors leads to enhanced tolerance to microbial, fungal, nematode or viral diseases, pathogens and pests), page 11, lines 30-31 (identification of relation between two nucleic acid molecules through hybridization under stringent conditions), and on page 12, lines 1-4 (nucleic acids hybridize under stringent conditions to a probe under wash conditions of 0.2 x SSC, 0.1% SDS at 65° C).

Support for new claim 62 is provided by, for example, page 5, line 9, and page 16, line 9 through page 17, line 24.

Support for new claim 63 is provided on, for example, page 3, lines 24-26, on page 13, lines 13-19, and page 14, lines 13-18.

Support for new claim 64 is provided on page 13, lines 15-19.

Support for new claim 65 is provided on page 13, lines 16-17 and page 13, line 20 through page 14, line 12.

Support for new claim 66 is provided on, for example, page 17, lines 10-11 (increasing tolerance or resistance to pathogens and pests), page 6, lines 29-31 and page 17, lines 10-23 (expression of instant AP2 transcription factors leads to enhanced tolerance to, fungal, and other pathogens), on page 19, lines 32 through page 21, line 28 (those genes induced after exposure to biotrophic fungal pathogens), and on page 24, line 9 to page 25, line 13 (analysis of *Arabidopsis* T2 progeny plants for pathogen resistance or pathogen tolerance).

Support for new claim 67 is provided on page 7, lines 22-28, page 17, lines 10-17, page 19, line 31 through page 21, line 28, page 24, line 9 through page 25, line 13. Statements in the

specification indicating that G28 can be used to confer tolerance to fungal pathogens have been confirmed by experimentation (see description in the Heard Declaration, above and below).

Support for new claim 68 is may be found in claim 5 as filed (method for altering the disease tolerance or resistance of a plant), page 6, lines 29-31 and page 17, lines 10-23 (expression of instant AP2 transcription factors leads to enhanced tolerance to microbial, fungal, nematode or viral diseases, pathogens and pests), page 11, lines 30-31 (identification of relation between two nucleic acid molecules through hybridization under stringent conditions), and on page 12, lines 1-4 (nucleic acids hybridize under stringent conditions to a probe under wash conditions of 0.2 x SSC, 0.1% SDS at 65° C).

Support for new claim 69 is provided by, for example, page 5, line 9, page 6, lines 28-31, and page 16, line 9 through page 17, line 24.

Support for new claim 70 is provided on, for example, page 3, lines 24-26, on page 13, lines 13-19, and page 14, lines 13-18.

Support for new claim 71 is provided on page 13, lines 15-19.

Support for new claim 72 is provided on page 13, lines 16-17 and page 13, line 20 through page 14, line 12.

Support for new claim 73 is provided by claim 5 as filed (method for altering the disease tolerance or resistance of a plant) and on, for example, page 17, lines 10-11 (increasing tolerance or resistance to pathogens and pests), page 6, lines 29-31 and page 17, lines 10-23 (expression of instant AP2 transcription factors leads to enhanced tolerance to, fungal, and other pathogens), on page 19, lines 32 through page 21, line 28 (those genes induced after exposure to biotrophic fungal pathogens), and on page 24, line 9 to page 25, line 13 (analysis of *Arabidopsis* T2 progeny plants for pathogen resistance or pathogen tolerance).

Support for new claim 74 is provided on page 7, lines 22-28, page 17, lines 10-17, page 19, line 31 through page 21, line 28, page 24, line 9 through page 25, line 13. Statements in the specification indicating that G28 can be used to confer tolerance to fungal pathogens have been confirmed by experimentation (see description in the Heard Declaration, above and below).

Support for new claim 75 may be found at, for example, page 17, lines 10-11 (utility for increasing tolerance or resistance to pathogens and pests), page 6, lines 29-31 and page 17, lines

10-23 (expression of instant AP2 transcription factors leads to enhanced tolerance to, fungal, and other pathogens), on page 19, line 32 through page 21, line 28 (those genes induced after exposure to biotrophic fungal pathogens), and on page 24, line 9 to page 25, line 13 (analysis of *Arabidopsis* T2 progeny plants for pathogen resistance or pathogen tolerance). The specification at page 10, lines 13-14 also notes conservative substitutions, deletions, and insertions of which one of skill in the art would be aware. Further support for residues that may be substituted and with which residues they may be substituted is disclosed in Table 1 at page 46 of provisional Application serial No. 60/125,814, from which the instant application claims priority.

Support for new claim 76 is provided on page 7, lines 22-28, page 17, lines 10-17, page 19, line 31 through page 21, line 28, page 24, line 9 through page 25, line 13. Statements in the specification indicating that G28 can be used to confer tolerance to fungal pathogens have been confirmed by experimentation (see description in the Heard Declaration, above and below).

RESPONSE TO REJECTIONS

Response to Rejection under 35 U.S.C. § 112, second paragraph

Claims 17-36 were rejected under 35 U.S.C. § 112, second paragraph, as they allegedly failed to particularly point out and distinctly claim the subject matter regarded as the invention. Applicants respectfully disagree.

Each of claims 17-36 have been canceled without prejudice or disclaimer and not for reasons of patentability. However, applicants submit that one of skill in the art would understand the metes and bounds of these claims. For example, the “conserved domain” of a transcription factor is a well-known concept in the field. Furthermore, the specification and the priority document specifically lists the conserved domains for a large number of transcription factors (see Table 1). Furthermore, that the conserved domains of transcription factors can be localization domains, activation domains, and DNA binding domains, for example, is also well known in the art. Applicants included a number of citations to references discussing the transcription factor families (see page 8, for example), and these and other references demonstrate that one of skill in

the art would understand the metes and bounds of each of the terms noted by the Examiner at page 5 of Paper No. 19.

The new claims 37-76 for the presently claimed transcription factors refer to only SEQ ID NO: 17 or 18 and are directed to a specific, defined conserved domain of residues 145-213 of SEQ ID NO: 18 which is supported by the specification as filed (see Figure 1a). That the conserved domains of transcription factors can be localization domains, activation domains, and DNA binding domains, for example, is also well known in the art. Applicants included a number of citations to references discussing the transcription factor families (*see* page 8, for example), and these and other references demonstrate that one of skill in the art would understand the metes and bounds of each of the terms noted by the Examiner at page 5 of Paper No. 19.

Applicants also submit that it would be abundantly clear to one of skill in the art what is meant by the term "conserved domain" in the present context. For example, the "conserved domain" of a transcription factor is a well-known concept in the field. See for example, Riechmann and Meyerowitz (1998), which is incorporated by reference (page 7, line 37 to page 8, line 1, and page 26, lines 29-30; submitted by Applicants in Information Disclosure Statement 11th October, 2002), and which defines conserved AP2 domains, with reference to, for example, the "conserved AP2 DNA-binding domain" (page 634, column 2, line 6); and "the conserved AP2 domain" (page 641, column 2, line 13). One of skill in the art would also be aware that the term "conserved domain" is often used in the art. For example, a search of issued U.S. patents yielded a sizeable number of references that make use of the term "conserved domain" (from 1996-2002 a search of ("conserved domain" AND sequence) revealed 338 patents that use the term in the present context).

The claims that were directed to "freezing, or nutrient, or pathogen stress" have been canceled, and the new claims are directed to "disease tolerance or resistance" or "tolerance to fungal disease", which are art-recognized terms and are defined in the specification; for example, page 7, lines 12-18, states: "[o]f particular interest are traits relating to increased disease resistance or tolerance of a plant, such as alterations in cell wall composition, trichome number or structure, callose induction, phytoalexin induction, alterations in the cell death response or the

like. These transgenic plants may be more resistant to biotrophic or necrotrophic pathogens such as a fungus, bacterium, mollicute, virus, nematode, a parasitic higher plant or the like and associated diseases. Another desirable phenotype is a change in the overall gene expression pattern of the plant in response to disease"). Page 17, lines 14-23 list pathogens and pests to which the plants can become tolerant.

Regarding the rejection based on improper Markush format, Applicant's believe the new claims avoid the rejection. The same may be said for the rejection based on indefiniteness of claims 22, 27 and 33, which referred to amino acid sequences, and claims that were dependent on indefinite claims.

The new claims make no reference to "DNA-binding domains." Therefore, even though one of skill in the art would clearly understand this term and its meaning, applicants need not address this reason for the rejection here.

Applicants submit that the new claims satisfy the requirements of 35 U.S.C. § 112, second paragraph, and respectfully request that this rejection be withdrawn.

Response to Rejection under 35 U.S.C. § 112, first paragraph

Claims 17-36 were rejected under 35 U.S.C. § 112, first paragraph, as the specification allegedly fails to reasonably convey that Applicants had possession of the invention. Applicants respectfully disagree with this rejection.

At page 6 of Paper No. 19, the Examiner characterizes the transcription factor of SEQ ID NO: 17 and 18 as "putative." Applicants respectfully but strongly disagree, as all the available and pertinent evidence indicates that SEQ ID NO: 17 encodes, and SEQ ID NO: 18 represents, a transcription factor, for at the very least the following reasons.

Applicants respectfully submit that the Examiner may not fully appreciate the PTO Written Description Examination Guidelines as they apply here. The USPTO Written Description Examination Guidelines state that:

(t)he written description requirement for a claimed genus may satisfied through sufficient description of a representative number of species by actual reduction to practice ..., reduction to drawings ..., or by disclosure of relevant, identifying characteristics, i.e.,

structure or other physical and/or chemical properties, by functional characteristics coupled with a known or disclosed correlation between function and structure, or by a combination of such identifying characteristics, sufficient to show the applicant was in possession of the claimed genus. (Federal Register, Jan. 5, 2001, Vol. 66(4) page 1106, II § 3 (a)(2).)

As shown in the attached report (Appendix F) and the Heard Declaration (Appendix C), Applicants correctly predicted the function of the sequence G28 (SEQ ID NO: 17; encoding SEQ ID NO: 18) based upon the identification of the sequence as an AP2 transcription factor and upon results from microarray experiments as disclosed in the instant application, and therefore Applicants characterized the sequence (G28) structurally and functionally. This is reflected by the present specification, which asserts in no uncertain terms that G28 is a transcription factor. See for example, page 4, line 30 through page 5, line 10: The transcription factor sequence may comprise a whole coding sequence or a fragment or domain of a coding sequence. A "fragment or domain, as referred to polypeptides, may be a portion of a polypeptide which performs at least one biological function of the intact polypeptide in substantially the same manner or to a similar extent as does the intact polypeptide. ... Exemplary polynucleotides or polypeptides comprise a sequence provided in the Sequence Listing as ... SEQ ID No.17 (G28), SEQ ID No.18 (G28 protein)." Applicants also clearly assert that G28 is a transcription factor in the specification at Figure 1a, because they assign a transcription factor family name (AP2) to the sequence.

Furthermore, as Fujimoto *et al.* ((2000) Plant Cell 12: 393-404, attached) point out, G28 (AtERF1; GenBank accession number AB008103) and related molecules are "members of a novel family of transcription factors that are specific to plants" (Abstract, lines 1-2). AtERF1 expression was also induced by ethylene and wounding, but not by other abiotic stresses such as cold, salinity, or drought (Fujimoto *et al.*, (2000), *supra*). Biotic stresses were apparently not tested. Since Fujimoto *et al.* published data identifying G28 as a transcription factor, they provide further evidence that Applicants were correct when they identified G28 as a transcription factor.

Furthermore, as Riechmann and Meyerowitz (1998) *supra*, point out "[f]ive amino acid residues are absolutely conserved among all of the AP2 domain sequences...[t]he conservation of those residues suggests their importance for the structure/function of the AP2 domain") p. 635,

column 2, lines 41-47). The sequence of G28 (SEQ ID NO: 18) with the five absolutely conserved residues shown in boldface in the underlined conserved domain below, are:

MSMTADSQSDYAFLESIRRHLLGESEPIlseSTASSVTQSCVTGQSIKPVYGRNPSFSKLYPCFTESWGDL
PLKENDESDMLVYGILNDAFHGGWEPSSSSSEDEDRSSFPSVKIETPESFAAVDSVPVKKEKTSFVSAAVTA
AKGKH~~YR~~**G**VRQRPW**G**KFAAEIRDPAKNGARVWLGTFTETAEDAALAY**D**RAAFRMR**G**SRALL**N**FPLRVNSG
EPDPVRIKSKRSSFSNNENGAPKKRRTVAAGGGMDKGLTVKCEVVEVARGDRLLVL

As noted in the Riechmann reference on page 635, column 2, lines 3-7, "[a] large number of sequences encoding AP2/EREBP proteins are already present in the databases and can be *easily identified* through BLAST searches due to the conservation of the AP2 domain, which, as mentioned above, is the only region conserved among all the proteins of the family (*emphasis added*).\" Thus, one of skill in the art would not doubt that G28, having an AP2 conserved domain, and possessing the prerequisite five amino acid residues that are conserved among all of the AP2 domain sequences, is a member of the AP2 family of transcription factors.

Finally, and perhaps most significantly, Applicants have demonstrated the function of G28 as a transcription factor. In the declaration under 37 C.F.R. § 1.132 from Dr. Jacqueline Heard (Appendix C), who is a listed inventor of the instant application, Dr. Heard submits results of experiments performed prior to the date of filing the instant application. The results are presented as tables which show that Applicants had identified numerous polynucleotide sequences which endogenous mRNA levels were up-regulated by pathogens and/or methyl jasmonate, including G28 (SEQ ID NO: 17), G4 (SEQ ID NO: 12), G19 (SEQ ID NO: 22), G1006 (SEQ ID NO: 34), and G25 (SEQ ID NO: 82). Dr Heard also submits an internal report which discloses the effect of overexpressing G28 (SEQ ID NO: 17) in a transgenic plant. The report shows that Applicants' SEQ ID NO: 17 (and hence the encoded SEQ ID NO: 18) confers tolerance to fungal pathogens, just as the G19 (SEQ ID NO: 22) sequence noted above does.

Thus, the knowledge in the art, the above-described disclosures from the specification defining G28 as a transcription factor, coupled with the data submitted, all point to the same conclusion: G28 is a transcription factor.

With regard to the statement in Paper 19 that "Applicant's definition of a 'conserved domain' appears to be relative to SEQ ID NO: 18," Applicants note that the new claims are now directed to transcription factors comprising a conserved domain of a plant AP2 transcription factor wherein the conserved domain comprises one of an amino acid sequence of residues 145-213 of SEQ ID NO: 18, or an amino acid sequence having at least 84% identity to residues 145-213 of SEQ ID NO: 18, or an amino acid sequence of residues 145-213 of SEQ ID NO: 18 having one or more conservative substitutions, deletions, or insertions.

Applicants clearly identified amino acid residues 145-213 of SEQ ID NO: 18 at Figure 1a. The specification on page 11, lines 21-22 points out that "transcription factors that are homologs of the disclosed sequences will typically share at least 40% amino acid sequence identity," and page 6, lines 4-6: "[a] 'conserved domain' refers to a polynucleotide or polypeptide fragment that is more conserved at a sequence level than other fragments when the polynucleotide or polypeptide is compared with homologous genes or proteins from other plants."

The specification at page 10, lines 13-14 also notes the conservative substitutions, deletions, and insertions of which one of skill in the art would be aware. Support for which residues may be substituted and which residues they may be substituted with, is further disclosed in Table 1 at page 46 of provisional Application serial No. 60/125,814 (Appendix D), from which the instant application claims priority. In addition, fragments of domains of transcription factor sequences are specifically disclosed at page 4, lines 32-37, and in Figure 1.

Applicants also note that SEQ ID NO: 22 (G19) is disclosed in the specification as being significantly induced upon infection by the fungal pathogen *Erysiphe orontii* (see specification at page 25, lines 4-9). As shown in the accompanying BLAST report (Appendix F), G19 and G22 share 42% identity (note that the specification discloses that transcription factors that are homologs will typically share at least 40% amino acid sequence identity (page 11, lines 21-22)). The conserved domain of G28 (SEQ ID NO: 18) shows even greater sequence identity to G19 (SEQ ID NO: 22), namely 66% identity (see attached BLAST report, page 11; handwritten notes regarding % identity). Thus, the BLAST report supports the new claims directed to an amino acid sequence having at least 42% identity to SEQ ID NO: 18. The use of identity ranges is well

known in the art and is specifically supported by the specification at, for example, page 11, lines 21-23. Thus, one of skill in the art would recognize the ability of the G19 sequence (SEQ ID NO: 22) and the closely related G28 sequence (SEQ ID NO: 18) to produce transgenic plants with disease tolerance or resistance, as is documented by Dr. Heard (see attached Declaration under 37 CFR 1.132, paragraphs 8 and 9 in particular).

The BLAST report also shows that conserved domains of G4 (SEQ ID NO: 12) with 62% identity to SEQ ID NO: 18, and G1006 (SEQ ID NO: 34) with 97% identity to SEQ ID NO: 18, and G25 (SEQ ID NO: 82) with 69% identity, show substantial sequence identity to the conserved domain of G28 (SEQ ID NO: 18) (*see* handwritten notes regarding % identity provided by the Applicants, particularly at pages 4, 10, and 13, respectively). The BLAST analysis method is fully disclosed in the specification at page 11, lines 3-20; at page 19, lines 2-8; and at page 25, lines 32-37. In addition, the specification discloses the function and utility of these polynucleotide and polypeptide sequences (see specification at pages 1 through 3, and pages 5 through 8, where the different polynucleotide and polypeptide sequences and their use are disclosed). Accordingly, Applicants' specification, in light of this evidence, clearly demonstrates the range of polynucleotide sequences that could be used in the new methods and transgenic plants.

As one of ordinary skill in the art recognizes, conserved domains may be identified as regions or domains of identity to a specific consensus sequence (see, for example, Riechmann et al., (2000) *Science* 290: 2105-2110; reference attached). Thus, by using alignment methods well known in the art, the conserved domains of the AP2 (APETALA2) domain transcription factor, the function of the presently claimed transcription factors may be determined family. Please see, for example, Riechmann and Meyerowitz (1998) *Biol. Chem.* 379: 633-646, which is incorporated by reference at page 7, line 37, continuing to page 8, line 1 of the specification (submitted by Applicants in Information Disclosure Statement 11th October, 2002). Riechmann and Meyerowitz specifically note:

(1) "Five amino acid residues are absolutely conserved among all of the AP2 domain sequences...[t]he conservation of those residues suggests their importance for the structure/function of the AP2 domain") p. 635, column 2, lines 41-47);

(2) "[a] primary source of specificity will be found to have very related, if not identical, DNA-binding specificities, similar to what has been observed for other large families of transcription factors" p. 643, column 1, lines 12-16; and

(4) "[p]artial genetic redundancy for AP2 function and/or postranscriptional mechanisms of control of AP2 activity may underlie" (p. 640, column 2, lines 32-35).

Thus, at the time the present application was filed, one of skill in the art would recognize that portions of the AP2 domain are *absolutely conserved*, which strongly suggests their importance in imparting function and functional overlap, and that AP2 transcription factors are distinguishable by differential DNA-binding specificities, which are likely to be related, if not identical, within the family, and that there is genetic redundancy for AP2 function. One of skill in the art would instantly recognize that data obtained with G19-overexpressing plants would lead to the prediction that similar results would be obtained with plants overexpressing a transcription factor gene with a highly similar AP2 domain, as indicated in the present application.

Based on these published assertions, one of skill in the art would clearly have recognized from the knowledge available in the art, and from Applicants' specification and Figures, that G28 is a transcription factor and has the "only region conserved among all the proteins of the family" (see reference to Riechmann (1998), above), where conserved domains are listed for numerous sequences including SEQ ID NO: 18. Accordingly, Applicants' adequately demonstrated that they possessed sequences comprising a conserved domain as claimed. One of skill in the art could clearly perform the BLAST searches in order to print out and physically possess the sequences having at least 84% identity to the conserved domain of SEQ ID NO: 18, as recited in the new claims. In view of these facts, the art-recognized use of the terms "transcription factor" and "conserved domain," and the definition of a conserved AP2 domain (see above), Applicants submit that the new claims satisfy the written description requirement of 35 U.S.C. § 112, first paragraph.

To further demonstrate that the functional characteristics recited in the new claims would have been understood to be present with the recited polynucleotides, or easily identified in

transgenic plants comprising these polynucleotides, Applicants also enclose a copy of Table 1 of U.S. provisional application serial no. 60/125,814, at page 46 (Appendix D). The present application claims priority to this provisional application. Table 1 shows amino acid residues of a polypeptide sequence that may be substituted with other amino acids and are restricted to conservative substitutions.

Accordingly, one of skill in the art would not doubt that Applicants have adequately described the recited polynucleotides, specified a novel and useful transcription factor function of G28 and other AP2 family transcription factors in plants, and demonstrated that they can be made and used to enhance plant disease tolerance, for example, or that transgenic plants as claimed can be made and used. Applicants respectfully request reconsideration and withdrawal of the rejection.

Accordingly, applicants submit that the new claims satisfy the written description requirement of 35 U.S.C. § 112, first paragraph.

Claims 17-36 were further rejected under 35 U.S.C. § 112, first paragraph, as the specification allegedly failed to describe the subject matter of the claims in such a way as to enable one of skill in the art to make and use it. Applicants respectfully disagree.

The Examiner asserts that Applicants' specification fails to teach how SEQ ID NO: 18 can enhance a plant tolerance to stresses from pathogens, disease, freezing, etc. (*see* page 8 of Paper No: 19). The new claims are not directed to disease and freezing stress. As noted above, Dr. Heard has submitted that Applicants have in their possession data showing that SEQ ID NO: 17 confers pathogen or disease tolerance (*see* attached Declaration Appendix C, at exhibit H). One of skill in the art would not doubt these data. Applicants respectfully request reconsideration.

The Examiner has also asserted that the Quattrocchio document implies that undue experimentation would be required to produce a desired phenotype (*see* page 9 of Paper No. 19). First, the experiments disclosed in the specification at page 20, lines 11-16, and at page 24, lines 11-37 and continued on page 25, lines 1-2, show how one of skill in the art can determine the

plant traits as recited in the claims. These are the same types of experiments used to produce the results in the report for the G28 sequence.

Second, a careful reading of Quattrocchio shows that it does not support the undue experimentation position. Quattrocchio discusses the *Petunia an2* gene, encoding a MYB-domain containing protein and that it had been suggested it could be the ortholog of the *Zea mays c1* gene (Quattrocchio, page 476, first column, second paragraph, lines 2-5). The *an2* gene had previously been shown to regulate anthocyanin biosynthesis (*see* Quattrocchio, page 478, first column, lines 1-2). Quattrocchio further states that “[t]he evolutionary model *predicts* that expression of *c1* and an *r* gene in *P. hybrida* induces *dfrA* but not *chsA*, *which was confirmed* ...” and that “[t]he model also *predicts* that AN2 (the protein encoded by the *an2* gene) should activate both early and late anthocyanin genes in *Z. mays*. *This is precisely what we observed experimentally*” (*see* Quattrocchio, page 486, first column, first paragraph, lines 5-10; *emphasis added*). Thus, Quattrocchio shows that the MYB-encoding polynucleotide/genes *an2* and *c1* are functionally interchangeable between different plant species and that the same pathway of genes is controlled by the orthologous MYB-domain containing protein.

This interchangeability is further supported by Quattrocchio’s statement that “[o]ur data for *P. hybrida* show that neither the *Z. mays* regulators *c1* and *lc* ... nor *an2* and *jaf13* from *P. hybrida* can activate the early flavonoid genes” Therefore, the orthologous MYB-encoding polynucleotide sequences do not activate transcription inappropriately under the experimental conditions observed (*see* Quattrocchio, page 486, first column, fourth paragraph, lines 3-5 and continued in second column, line 1).

Nowhere does Quattrocchio suggest that the effect of transforming a plant with a heterologous “MYB” encoding polynucleotide “is unpredictable.” In fact, Quattrocchio strongly suggests the very opposite, that transforming a plant with a heterologous MYB-encoding polynucleotide results in a predictable effect on the **same** pathway.

Also, Dr. Heard addresses the understanding one of skill in the art has for the contents of Quattrocchio in her declaration (*see* page 4, in particular). Dr. Heard explains that the Quattrocchio document actually supports the predictable nature of using transcription factors to produce predictable effects on the same traits in plants. Further, the Duggleby (of record) and

Quattrocchio documents together, as discussed at pages 3-4 of the declaration, provide additional evidence of the ability of one of skill in the art to use the sequences recited in the claims. These statements demonstrate that the conclusions the Patent Office asserts from the Quattrocchio document are incorrect.

Accordingly, Applicants respectfully request reconsideration of the Quattrocchio document and this rejection.

Applicants also submit that a *prima facie* case of lack of enablement has not been made.

The remaining arguments are directed to the "homologous window sequence" recitation. Applicants note that the new claims do not recite the "homologous window sequence" discussed in this rejection. Accordingly, Applicants respectfully request reconsideration and that the Examiner withdraw the rejection.

Response to Rejection under 35 U.S.C. § 102

Claims 17-36 were rejected under 35 U.S.C. § 102(b) as being anticipated by Martin *et al.* (WO 97/47183).

Applicants note that the new claims do not recite a "homologous window sequence" or the "at least 6 consecutive amino acids" discussed in this rejection. Furthermore, the SEQ ID NO: 4 of Martin does not display the sequence identity to amino acid residues 145-213 of SEQ ID NO: 18 now recited in the claims. Accordingly, Applicants believe there is no anticipation of the present claims by the Martin reference.

Applicants respectfully request reconsideration and withdrawal of this rejection.

Provisional Double Patenting Rejection

The Examiner has rejected Claims 17-36 under the judicially created doctrine of obviousness-type double patenting as being unpatentable over copending applications 09/713,994 and 09/934,455.

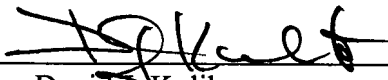
As this is a provisional rejection, Applicants will address this rejection when claims of this or the copending application are allowed.

Applicants have requested an extension of time of two months in order to respond to the outstanding Office Action. No additional extension of time fees or requests for extension of time, or any other fees or petitions, are believed to be necessary to enter and consider this paper. If, however, any petitions or extensions of time are required or any fees are due in order to enter or consider this paper or enter or consider any paper accompanying this paper, including fees for net addition of claims, or in order to keep this application pending, Applicants hereby request any extensions or petitions necessary and the Commissioner is hereby authorized to charge Deposit Account No. 50-1129 for any fees.

Respectfully submitted,

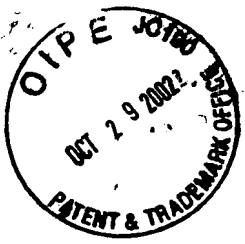
WILEY REIN & FIELDING LLP

Date: October 29, 2002

By: 
David J. Kulik
Registration No. 36,576

Enclosures: Appendix A (Marked-up version of amended Specification pages – 2 pages)
Appendix B (Gene Expression data; G1006 and G25 – 1 page)
Appendix C (executed declaration of Dr. Heard, pages 1-5 and Exhibits A-G)
Appendix D (pages 45 and 46 of priority application 60/125,814)
Appendix E (2 page conserved domain sequence comparison data for G28, G4, G19, G25, and G1006)
Appendix F (14 page BLAST report for G28, with handwritten notes)\
and
Fujimoto *et al.*, The Plant Cell 12:393-404 (2000)
Riechmann *et al.*, Science 290:2105-2110 (2000)

Wiley Rein & Fielding LLP
Attn: Patent Administration
1776 K Street, N.W.
Washington, D.C. 20006
202.719.7000 (Telephone)
202.719.7049 (Facsimile)



APPENDIX A



APPENDIX A

Marked-up Version of Amended Specification at pages 7-8

These transcription factors can be used to modulate a plant's response to disease. The plant transcription factors may belong to one of the following transcription factor families: the AP2 (APETALA2) domain transcription factor family (Riechmann and Meyerowitz (1998) [J.] Biol. Chem. 379:633-646); the MYB transcription factor family (Martin and Paz-Ares, (1997) Trends Genet. 13:67-73); the MADS domain transcription factor family (Riechmann and Meyerowitz (1997) [J.] Biol. Chem. 378:1079-1101); the WRKY protein family (Ishiguro and Nakamura (1994) Mol. Gen. Genet. 244:563-571); the ankyrin-repeat protein family (Zhang et al. (1992) Plant Cell 4:1575-1588); the zinc finger protein (Z) family (Klug and Schwabe (1995) FASEB J. 9: 597-604); the homeobox (HB) protein family (Buerklin, in: Duboule (1994) Guidebook to the Homeobox Genes, Oxford University Press, pp. 27-71); the CAAT-element binding proteins (Forsburg and Guarente (1989) Genes Dev. 3:1166-1178); the squamosa promoter binding proteins (SPB) (Klein et al. (1996) Mol. Gen. Genet. 1996 250:7-16); the NAM protein family (Souer et al. (1996) Cell 85:159-170); the IAA/AUX proteins (Rouse et al. (1998) Science 279:1371-1373); the HLH/MYC protein family (Littlewood et al. (1994) Prot. Profile 1:639-709); the DNA-binding protein (DBP) family (Tucker et al. (1994) EMBO J. 13:2994-3002); the bZIP family of transcription factors (Foster et al. (1994) FASEB J. 8:192-200); the Box P-binding protein (the BPF-1) family (da Costa e Silva et al. (1993) Plant J. 4:125-135); the high mobility group (HMG) family (Bustin and Reeves (1996) Prog. Nucl. Acids Res. Mol. Biol. 54:35-100); the scarecrow (SCR) family (Di Laurenzio et al. (1996) Cell 86:423-433); the GF14 family (Wu et al. (1997) Plant Physiol. 114:1421-1431); the polycomb (PCOMB) family (Kennison (1995) Annu. Rev. Genet. 29:289-303); the teosinte branched (TEO) family (Luo et al. (1996) Nature 383:794-799); the ABI3 family (Giraudat et al. (1992) Plant Cell 4:1251-1261); the triple helix (TH) family (Dehesh et al. (1990) Science 250:1397-1399); the EIL family (Chao et al. (1997) Cell 89:1133-44); the AT-HOOK family (Reeves and Nissen (1990) Journal of Biological Chemistry 265:8573-8582); the SIFA family (Zhou et al. (1995) Nucleic Acids Res. 23:1165-1169); the bZIPT2 family (Lu and Ferl (1995) Plant Physiol. 109:723); the YABBY

family (Bowman et al. (1999) Development 126:2387-96); the PAZ family (Bohmert et al. (1998) EMBO J. 17:170-80); a family of miscellaneous (MISC) transcription factors including the DPBF family (Kim et al. (1997) Plant J. 11:1237-1251) and the SPF1 family (Ishiguro and Nakamura (1994) Mol. Gen. Genet. 244:563-571); the golden (GLD) family (Hall et al. (1998) Plant Cell 10:925-936).

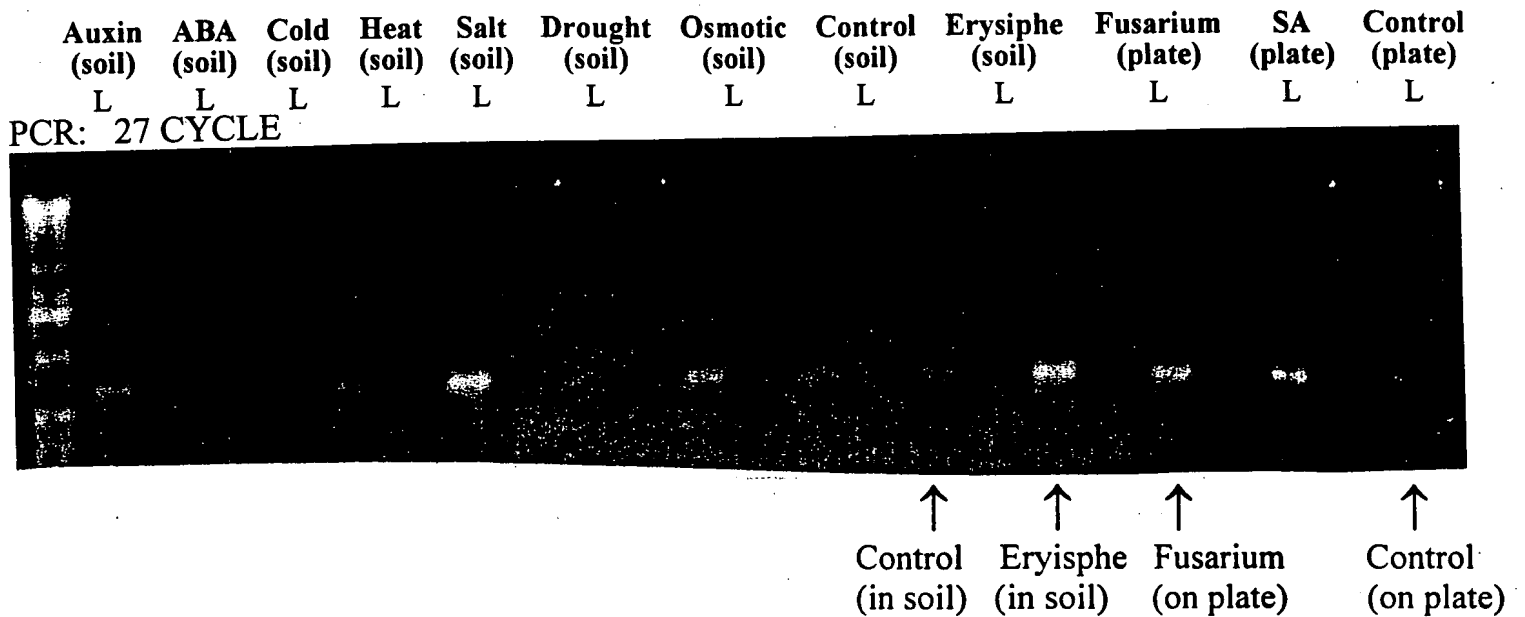
WRFMAIN 1184985.1

APPENDIX B

Appendix B

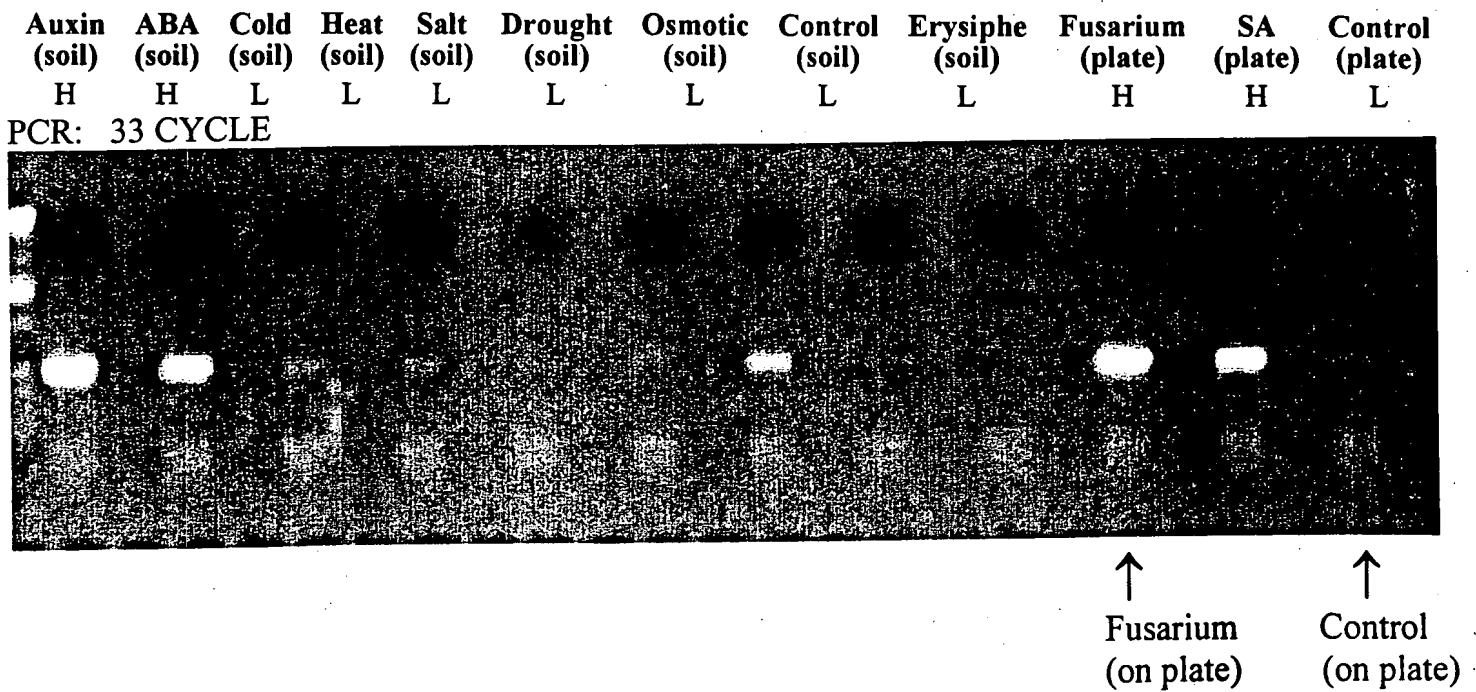
G1006

Gene Expression in Different Conditions



G25

Gene Expression in Different Conditions





APPENDIX C

Executed Declaration of Jacqueline E. Heard Under 37 CFR 1.132

Exhibit A: *Fusarium* treatment after 24 hours

Exhibit B: *Fusarium* treatment after 48 hours

Exhibit C: *Erysiphe* treatment after 7 days

Exhibit D: Methyl Jasmonate treatment after 24 hours

Exhibit E: *Botrytis* treatment after 12 hours

Exhibit F: *Erysiphe* treatment

Exhibit G: Methyl Jasmonate treatment

Exhibit H: Summary of Overexpressor G28, Family AP2

APPENDIX D

share at least 60%, 90% or 95% sequence identity. The homologs will also have substantially the same DNA binding specificity. At the nucleotide level, the sequences will typically share at least 40% nucleotide sequence identity, preferably at least 50%, 60%, 70% or 80% sequence identity, and more preferably 85%, 90%, 95% or 97% sequence identity.

Homologs from the same plant, different plant species or other organisms may be identified using sequence alignment methods and homology calculations, such as those described in Altschul et al. (1994) *Nature Genetics* 6: 119-129. For example, the NCBI Basic Local Alignment Search Tool (BLAST) (Altschul et al. (1990) Basic local alignment search tool. *J. Mol. Biol.* 215:403-410). is available from several sources, including the National Center for Biotechnology Information (NCBI, Bethesda, MD, for use in connection with the sequence analysis programs blastp, blastn, blastx, tblastp, tblastn and tblastx.

Substitutions, deletions and insertions introduced in the DNA binding domain are also envisioned by this invention. Amino acid substitutions are typically of single residues; insertions usually will be on the order of about from 1 to 10 amino acid residues; and deletions will range about from 1 to 30 residues. Deletions or insertions preferably are made in adjacent pairs, i.e., a deletion of two residues or insertion of two residues. Substitutions, deletions, insertions or any combination thereof may be combined to arrive at a final construct. Obviously, the mutations that are made in the DNA encoding the protein must not place the sequence out of reading frame and preferably will not create complementary regions that could produce secondary mRNA structure.

Substitutional variants are those in which at least one residue in the amino acid sequence has been removed and a different residue inserted in its place. Such substitutions generally are made in accordance with the following Table 1 when it is desired to finely modulate the characteristics of the protein. Table 1 shows amino acids which may be substituted for an original amino acid in a protein and which are typically regarded as conservative substitutions.

Table 1

Original Residue	Conservative Substitutions
Ala	ser
Arg	lys
Asn	gln; his
Asp	glu
Cys	ser
Gln	asn
Glu	asp
Gly	pro
His	asn; gln
Ile	leu, val
Leu	ile; val
Lys	arg; gln; glu
Met	leu; ile
Phe	met; leu; tyr
Ser	thr
Thr	ser
Trp	tyr
Tyr	trp; phe
Val	ile; leu

An alternative indication is to show whether two nucleic acid molecules are closely related is that the two molecules hybridize to each other under stringent conditions. Stringent conditions are sequence dependent and are different under different environmental parameters. Generally, stringent conditions are selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Conditions for nucleic acid hybridization and calculation of stringencies can be found in Sambrook et al., *Molecular Cloning. A Laboratory Manual*, Ed. 2, Cold Spring Harbor Laboratory Press, New York (1989)) and Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes* Part I, Chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier, New York (1993). Nucleic acid molecules that hybridize under stringent conditions will typically hybridize to a probe based on either the entire cDNA

APPENDIX E

Appendix E

Percent identity comparisons of conserved domains (underlined) of subject sequences of AP2 transcription factor family members as disclosed in specification USSN 09/533,029.

SEQ ID NO:18 (G28) (query sequence) Conserved domain = residues 145-213
MSMTADSQSDYAFLESIRRHLLGESEPIIESTASSVTQSCVTGQSIKPVYGRNPSFSKLYPCFTESWGDLPKEND
SEDMLVYGILNDAFHGGWEPSSSSSEDRSSFPSVKIETPESFAAVDSVPVKKEKTSVPVSAAVTAAKGKHYRGVRQR
PWGKFAAEIRDPKNGARVWLGTFTETAEDAALAYDRAAFMRGRSALLNFPLRVNSGEPDPVRIKSKRSSFSSSNEN
GAPKKRRTVAAGGGMDKGLTVKCEVVEVARGDRLLVL

BLAST comparisons of subject sequences disclosed in specification USSN 09/533,029

SEQ ID NO:12 (G4) Conserved domain = residues 121-188
MCGGAIISDFIPPPRSLRVTNEFIWPD LKNKVKASKKRSNKRSDFFDLDDDFEADFQGFKDDSAFDCEDDDVFN
KPFVFTATTKPVASAFVSTVGSAYAKKTVESAEQAEKSSKRKRKNQYRGIRQRPWGKWAAEIRDPKGSREWLGT
TAEAAARAYDAAARRIRGTKAKVNFPEEKNPVSVQKRPSAKTNNLQKSVAKPNKSVTLVQQPTHLSQQYCNSFDN
SFGDMSFMEEKPQMYNNQFGLTNSFDAGGNGYQYFSSDQGSNSFDCSEFGWSDHGPKTPEISSMLVNNNEASFVEE
TNAAKKLKPNSESDDDL MAYLDNALWDTPLEVEAMLGADAGAVTQEEENPVELWSLDEINFMLEGDF

SEQ ID NO:12 (G4) (Amino Acid Sequence) (gf=1)
Length = 375

Score = 90.5 bits (223), Expect = 1e-20
Identities = 43/62 (69%), Positives = 52/62 (83%), Gaps = 1/62 (1%)

Query: 4 YRGVRQRPWGKFAAEIRDPKNGARVWLGTFTETAEDAALAYDRAAFMRGRSALLNFPLR 63
YRG+RQRPWGK+AAEIRDP K G+R WLGT+TAE+AA AYD AA R+RG++A +NFP
Sbjct: 124 YRGIRQRPWGKWAAEIRDPK-GSREWLGTFTD TAEAAARAYDAAARRIRGTKAKVNFPEE 182

Query: 64 VN 65
N
Sbjct: 183 KN 184

SEQ ID NO:22 (G19) Conserved domain = residues 76-145
MCGGAIISDYAPLVTKAKGRKLTAEELWSEL DASAADDFWGFYSTSKLHPTNQVNVKEEAVKKEQATEPGKRRKRKN
VYRGIRKRPWGKWAAEIRDPKGVVWLGTFTNTAEAAAMAYDVAAKQIRGDKAKLNF PDLHHPPPNYTPPPSSPRS
TDQPPAKKVCVVSQSESELSQPSFPVECIGFGNGDEFQNL SYGFEPDYDLKQQISSLESFLELDGNTAEQPSQLDES
VSEVDMWMLDDVIASYE

SEQ ID NO:22 (G19) (Amino Acid Sequence) (gf=1)
Length = 248

Score = 89.7 bits (221), Expect = 2e-20
Identities = 42/58 (72%), Positives = 50/58 (85%), Gaps = 1/58 (1%)

Query: 4 YRGVRQRPWGKFAAEIRDPKNGARVWLGTFTETAEDAALAYDRAAFMRGRSALLNFP 61
YRG+R+RPWGK+AAEIRDP K G RVWLGT TAE+AA+AYD AA ++RG +A LNFP
Sbjct: 79 YRGIRKRPWGKWAAEIRDPK-GVRVWLGTFTNTAEAAAMAYDVAAKQIRGDKAKLNF 135

SEQ ID NO:82 (G25) Conserved domain = residues 47-114

MCGGAIISDFIWSKSESEPSQLGSVSSRKKRKPVSVSEERD GK RERK NLYRGIRQRPWGK WAAEIRDPSKGV RVWLGT
TFKTADEAARAYDVAAIKIRGRKAKLNFPNTQVEEEADTKPGGNQNELISENQVESLS E DLMAL E DYMRFYQIPVAD
DQSATDIGNLWSYQDSN

SEQ ID NO:82 (G25) (Amino Acid Sequence) (gf=1)

Length = 171

Score = 90.9 bits (224), Expect = 8e-21

Identities = 42/58 (72%), Positives = 51/58 (87%), Gaps = 1/58 (1%)

Query: 4 YRGVQRQRPWGKFAAEIRDPAKNGARVWLGTFFETAEDAALAYDRAAFMRGSRALLNFP 61
YRG+RQRPWGK+AAEIRDP+K G RVWLGTFF+TA++AA AYD AA ++RG +A LNFP
Sbjct: 50 YRGIRQRPWGK WAAEIRDPSK-GV RVWLGTFFKTADEAARAYDVAAIKIRGRKAKLNFP 106

SEQ ID NO:34 (G1006) Conserved domain = residues 114-182

MYGQCNIESDYALLESITRHLGGGGENELRLNESTPSSCFTESWGGLPLKENDSE DMLVYGLLKDAFHFDTS SSDL
SCLFD FPAVKVEPTENFTAMEEKPKKAIPVTETAVKAKHYRGVQRQRPWGKFAAEIRDPAKNGARVWLGTFFETAEDAA
LAYDIAAFMRGSRALLNFP LRVNSGEPDPVRITSKRSSSSSSSSSSSTSSSENGKLKRRRKAENLTSEVVQVKCEV
GDETRVDELLVS

SEQ ID NO:34 (G1006) (Amino Acid Sequence) (gf=1)

Length = 243

Score = 140 bits (353), Expect = 9e-36

Identities = 67/68 (98%), Positives = 67/68 (98%)

Query: 2 KHYRGVQRQRPWGKFAAEIRDPAKNGARVWLGTFFETAEDAALAYDRAAFMRGSRALLNFP 61
KHYRGVQRQRPWGKFAAEIRDPAKNGARVWLGTFFETAEDAALAYD AAFMRGSRALLNFP
Sbjct: 115 KHYRGVQRQRPWGKFAAEIRDPAKNGARVWLGTFFETAEDAALAYDIAAFMRGSRALLNFP 174

Query: 62 LRVNSGEP 69

LRVNSGEP

Sbjct: 175 LRVNSGEP 182

APPENDIX F

G28 (Amino Acid Sequence) (gf=1)	535
G1006 (Amino Acid Sequence) (gf=1)	268
G22 (Amino Acid Sequence) (gf=1)	2e-73
G1004 (Amino Acid Sequence) (gf=1)	161
G29 (Amino Acid Sequence) (gf=1)	2e-41
G1268 (Amino Acid Sequence) (gf=1)	6e-31
G46 (Amino Acid Sequence) (gf=1)	127
G1265 (Amino Acid Sequence) (gf=1)	117
G44 (Amino Acid Sequence) (gf=1)	6e-28
G45 (Amino Acid Sequence) (gf=1)	113
G47 (Amino Acid Sequence) (gf=1)	7e-27
G48 (Amino Acid Sequence) (gf=1)	110
G49 (Amino Acid Sequence) (gf=1)	4e-26
G50 (Amino Acid Sequence) (gf=1)	535

G16 (Amino Acid Sequence) (gf=1)
G2134 (Amino Acid Sequence) (gf=1)
G977 (Amino Acid Sequence) (gf=1)
G24 (Amino Acid Sequence) (gf=1)
G1387 (Amino Acid Sequence) (gf=1)
G1386 (Amino Acid Sequence) (gf=1)
G1379 (Amino Acid Sequence) (gf=1)
G441 (Amino Acid Sequence) (gf=1)
G39 (Amino Acid Sequence) (gf=1)
G2593 (Amino Acid Sequence) (gf=1)
G1946 (Amino Acid Sequence) (gf=1)
G1755 (Amino Acid Sequence) (gf=1)
G1277 (Amino Acid Sequence) (gf=1)
G21 (Amino Acid Sequence) (gf=1)
G2573 (Amino Acid Sequence) (gf=1)
G27 (Amino Acid Sequence) (gf=1)
G1381 (Amino Acid Sequence) (gf=1)
G2510 (Amino Acid Sequence) (gf=1)
G2291 (Amino Acid Sequence) (gf=1)
G2108 (Amino Acid Sequence) (gf=1)
G3 (Amino Acid Sequence) (gf=1)
G2115 (Amino Acid Sequence) (gf=1)
G869 (Amino Acid Sequence) (gf=1)
G40 (Amino Acid Sequence) (gf=1)
G41 (Amino Acid Sequence) (gf=1)
G2067 (Amino Acid Sequence) (gf=1)
G28 (Amino Acid Sequence) (gf=1)
Length = 268

Score = 535 bits (1378), Expect = e-154
Identities = 268/268 (100%), Positives = 268/268 (100%)

Query: 1 MSMTADSQSYAPLESIRRHLLGSEPIILSESTASSVTQSGIKPVYGRNPSPSKL 60
MSMTADSQSYAPLESIRRHLLGSEPIILSESTASSVTQSGIKPVYGRNPSPSKL
Sbjct: 1 MSMTADSQSYAPLESIRRHLLGSEPIILSESTASSVTQSGIKPVYGRNPSPSKL 60
Query: 61 YPCFTESWGDLPLKENSEDMLVYGLINDAFHGGWEPSSSSDEDRSPSVKIETPESP 120
YPCFTESWGDLPLKENSEDMLVYGLINDAFHGGWEPSSSSDEDRSPSVKIETPESP
Sbjct: 61 YPCFTESWGDLPLKENSEDMLVYGLINDAFHGGWEPSSSSDEDRSPSVKIETPESP 120
Query: 121 AAVDSVPVKKETSPVSAVTAAGKHGVRQRPWGKFAAIRDPAKNGARVWLGTPT 180
AAVDSVPVKKETSPVSAVTAAGKHGVRQRPWGKFAAIRDPAKNGARVWLGTPT
Sbjct: 121 AAVDSVPVKKETSPVSAVTAAGKHGVRQRPWGKFAAIRDPAKNGARVWLGTPT 180
Query: 181 AEDALAYDRAAFMRGSRALLNPLRVNSGEPDPVRIKSSPSSSNENGAPKKRTV 240
AEDALAYDRAAFMRGSRALLNPLRVNSGEPDPVRIKSSPSSSNENGAPKKRTV
Sbjct: 181 AEDALAYDRAAFMRGSRALLNPLRVNSGEPDPVRIKSSPSSSNENGAPKKRTV 240
Query: 241 AAGGMDKGLTVKCEVEVARGDRLLV 268
AAGGMDKGLTVKCEVEVARGDRLLV
Sbjct: 241 AAGGMDKGLTVKCEVEVARGDRLLV 268

>G1006 (Amino Acid Sequence) (gf=1)
Length = 243

Score = 268 bits (685), Expect = 2e-73
Identities = 163/271 (60%), Positives = 180/271 (66%), Gaps = 47/271 (17%)
Query: 8 QSDYAFLESIRRHLLG---ESEPIILSESTASSVTQSGIKPVYGRNPSPSKLYPCF 64
+SDYA LESI RHLIG E-E L-EET SS
Sbjct: 8 ESDYALLESITRHLGGGENELRLNESTPSS-----CF 41

Query: 65 TESWGDLPLKENSEDMLVYGLINDAFHGGWEPSSSSDEDRSPSVKIETPSPAAYD 124
TESWGDLPLKENSEDMLVYGLINDAFHGGWEPSSSSDEDRSPSVKIETPSPAAYD 124
Sbjct: 42 TESWGDLPLKENSEDMLVYGLINDAFHGGWEPSSSSDEDRSPSVKIETPSPAAYD 98
TESWGDLPLKENSEDMLVYGLINDAFHGGWEPSSSSDEDRSPSVKIETPSPAAYD 98
Query: 125 SVFVKKETSPVSAVTAAGKHGVRQRPWGKFAAIRDPAKNGARVWLGTPTFAEDA 184
SVFVKKETSPVSAVTAAGKHGVRQRPWGKFAAIRDPAKNGARVWLGTPTFAEDA 184
Sbjct: 99 EKP---KKAIPVTE---TAVKAKHVRQRPWGKFAAIRDPAKNGARVWLGTPTFAEDA 153
EKP---KKAIPVTE---TAVKAKHVRQRPWGKFAAIRDPAKNGARVWLGTPTFAEDA 153
Query: 185 ALAYDRAAFMRGSRALLNPLRVNSGEPDPVRIKSSPSSSSDEDRSPSVKIETPSPAAYD 236
ALAYDRAAFMRGSRALLNPLRVNSGEPDPVRIKSSPSSSSDEDRSPSVKIETPSPAAYD 236
Sbjct: 154 ALAYDRAAFMRGSRALLNPLRVNSGEPDPVRIKSSPSSSSDEDRSPSVKIETPSPAAYD 213
ALAYDRAAFMRGSRALLNPLRVNSGEPDPVRIKSSPSSSSDEDRSPSVKIETPSPAAYD 213
Query: 237 RRTAAGGMDKGLTVKCEVEVARGDRLLV 267
RRTAAGGMDKGLTVKCEVEVARGDRLLV 267
Sbjct: 214 RR--KAENLTSEVVQVKEVGEVDETRVDELLV 242
RR--KAENLTSEVVQVKEVGEVDETRVDELLV 242

>G22 (Amino Acid Sequence) (gf=1)
Length = 226

Score = 161 bits (408), Expect = 2e-41
Identities = 87/175 (49%), Positives = 107/175 (60%), Gaps = 21/175 (12%)
Query: 50 VYGRNPSPSKLYPCFTESWGDLPLKENSEDMLVYGLINDAFHGGWEPSSSSDEDRSP 109
VYGRNPSPSKLYPCFTESWGDLPLKENSEDMLVYGLINDAFHGGWEPSSSSDEDRSP 109
Sbjct: 15 MYRNPSPSNVI--LNDNMSDPLSLVDDSDQMAIYNTLRDAVSSGWTFS-----V 62
MYRNPSPSNVI--LNDNMSDPLSLVDDSDQMAIYNTLRDAVSSGWTFS-----V 62

Query: 110 PSVKIETPSPAAYDVPVKKETSPVSAVTAAGKHGVRQRPWGKFAAIRDPAKNGARVWL 169
PSVKIETPSPAAYDVPVKKETSPVSAVTAAGKHGVRQRPWGKFAAIRDPAKNGARVWL 169
Sbjct: 63 PVTSPAEEN-----KPPATKASGSHAPRQGMQYGRVRRPWPWGKFAAIRDPAKNGARVWL 113
PVTSPAEEN-----KPPATKASGSHAPRQGMQYGRVRRPWPWGKFAAIRDPAKNGARVWL 113

Query: 170 GARVWLGTPTFAEDAALAYDRAAFMRGSRALLNPLRVNSGEPDPVRIKSS 224
GARVWLGTPTFAEDAALAYDRAAFMRGSRALLNPLRVNSGEPDPVRIKSS 224
Sbjct: 114 GARVWLGTPTFAEDAALAYDRAAFMRGSRALLNPLRVNSGEPDPVRIKSS 168
GARVWLGTPTFAEDAALAYDRAAFMRGSRALLNPLRVNSGEPDPVRIKSS 168

>G1004 (Amino Acid Sequence) (gf=1)
Length = 300

Score = 127 bits (318), Expect = 6e-31
Identities = 91/227 (40%), Positives = 126/227 (55%), Gaps = 30/227 (13%)
Query: 17 IRRHLLGSEPIILSES---TASSVTQSCVTQSGIKPVYGRNPSPSKLYPC-FTESWGDL 71
IRRHLLGSEPIILSES---TASSVTQSCVTQSGIKPVYGRNPSPSKLYPC-FTESWGDL 71
Sbjct: 13 IEKHLLEASPVATDPMKHESSATSSSDSSSI--IFGS--SSSFAPIDFSESVCYP 68
IEKHLLEASPVATDPMKHESSATSSSDSSSI--IFGS--SSSFAPIDFSESVCYP 68

Query: 72 PLKENSEDMLVYGLINDAFHGGWEPSSSSDEDRSPSVKIETPSPAAYD 116
PLKENSEDMLVYGLINDAFHGGWEPSSSSDEDRSPSVKIETPSPAAYD 116
Sbjct: 69 EIIDLTPRSMEP--LSIPFEPDSEVSDFDPKPSNQNONQPEPELKSQIRKEPLKISL 126
EIIDLTPRSMEP--LSIPFEPDSEVSDFDPKPSNQNONQPEPELKSQIRKEPLKISL 126

Query: 117 PESPAANDVPVKKETSPVSAVTAAGKHGVRQRPWGKFAAIRDPAKNGARVWL 175
PESPAANDVPVKKETSPVSAVTAAGKHGVRQRPWGKFAAIRDPAKNGARVWL 175
Sbjct: 127 P---AKTEWIOFAAETKPEVTKEEKKHVRQRPWGKFAAIRDPAKNGARVWL 183
P---AKTEWIOFAAETKPEVTKEEKKHVRQRPWGKFAAIRDPAKNGARVWL 183

Query: 176 GTFETAEDALAYDRAAFMRGSRALLNPLRVNSGEPDPVRIKSS 222
GTFETAEDALAYDRAAFMRGSRALLNPLRVNSGEPDPVRIKSS 222
Sbjct: 184 GTFETAEDALAYDRAAFMRGSRALLNPLRVNSGEPDPVRIKSS 230
GTFETAEDALAYDRAAFMRGSRALLNPLRVNSGEPDPVRIKSS 230

>G29 (Amino Acid Sequence) (gf=1)
Length = 282

Score = 117 bits (292), Expect = 6e-28
Identities = 65/145 (44%), Positives = 84/145 (57%), Gaps = 6/145 (4%)

G28 (Length: 268)

Query: 88 NDAPHGWEPESSSSSDEDRSPFSVKIETPESFAAVDSVPVKKEKTSVPSAAVTAAGKH 147
ND+P P + +R P +KI P + +PV V A + +H

Sp1ct: 80 NDSEFTPOSPPRVTVOSNRK--PPLKIAPPNRKWIQFATGNPKPELPVV-VAAEEKRH 136

Query: 148 YGVQRQPPGKGAEEIRDPAKNGARVWLCTFETAEDAAALAYDRAAPRMGRSALLNPPLR 207
YGVQR RPWGKGAEEIRD P + G RVWLCTFETA + AA AYD+ APR+RGS+A+LNPLP
Spct: 137 YGVQMPGKGAEEIRDPTTRGTRVWLCTFETAIRAAAYDKAPFIRGSKAIIINPPLR 196

Query: 208 VNSGPPDPVRIKSKRSSPSSSNG 232
V+ P R + R ++ +G
Shift: 197 VTKWNP--RARDGGLYNKRKPDG 218

```
>Cl266 (Amino Acid Sequence) (gf=1)
Length = 218
```

Score = 113 bits (283), Expect = 7e-27
Identities = 61/134 (45%), Positives = 79/134 (58%), Gaps = 31/134 (23%)

Query: 71 LPLKENSEDLVYGIILNDAFHGGWPEPSSSSDEDRSPSVKIETPESFAAVDSVPVK 130
 LP ENDSE+M +YG++ + T +++ DS
 36 Y DENVNDGCPDMEI VCI TPDS TOGQVINDSDS 65

Query: 131 EKTSPVSAAVTAAKGHYRGVRQRPWGKPAABIRDPAKNGCARVWLGTPTAEADAALAYDR 190
+ P+ + K YRGVR+RPWGKFAAEIRD +NG RVNLGTFR+AE+AAALAYD+

Sbjct: 66 -QDLPIKSVSSNKKSEKSYRGVRRRPWGKPAAEIRDSRNGIRVWLGTFSBAEEAALAYDQ 124
 Query: 191 AAFMRGGRALLNF 204

AAF MRGS A+LNP
Sbjct: 125 AAFMRGSSA LNP 138

>G46 (Amino Acid Sequence) (gf=1)
Length = 207

Score = 110 bits (27%), Expect = 4e-26
Identities = 73/154 (47%), Positives = 89/154 (57%), Gaps = 16/154 (10%)

Query: 71 LPLKNDSEDMLVYGIINDAFHCGWEPSSSSSSDSDRSSFPVKIETPESFAAVDSVPVK 130
+P+E DS + +N+ E SSSS E SS S ET +S

Spict: 48 VPKORDSPVLDSDSFWRIQVEGRSSSSSSSPINSS--SSVPTDQV----- 95

Query: 131 EKTSPVSAAVTAAKGKHVGRVQRPMGKFAAEIRDPKNGARVWLGTFTAEADAALAYDR 190
+K V A +HYRGVR+RPMGKFAAEIRDPK G+R+WLGTFT+DA A AYD

Sbjct: 96 KKAERFEEVD- --RHYRGVRRRPWGKFAAEIRDPAKKGSRIWLGTFFSDVDAARAYDC 152
Query: 191 AAFMRGSRALLNFFPLRVNSGE-PDPVRIKSKRS 223

AAP++RG +A+INFLP E P K KRS
 Sbjct: 153 AAFKLGRKAVINFLDAGKYEPANSGRKRKRS 186

```
>G43 (Amino Acid Sequence) (gf=1)
      Length = 201
```

Score = 110 bits (276), Expect = 4e-26
Identities = 62/128 (48%), Positives = 82/128 (63%), Gaps = 3/128 (2%)

Query: 96 · EPSSSDEDRSSFPVKIETPESFAAVDSVPVKKEKTSVSAATAAKGKHGVRQRP 155
EPSS D D +++E S ++ + + E S + HYRGVR+RP
Sbjct: 55 EPSSVLDPPDSYVOEILOMEAESSSSSSTTTSPDEVTVSNRKKTRPRTRHYGVRPRP 114

Query: 156 WGKFAAEIRDPKNGARVWLGTFTAEADAALAYDRAAFPMRGSRALLNPLLRVNSGEPD- 214
WGKFAAEIRDPK G-R+WLGTFT+ DAA AYD AAR++RG +A+INEPL ++G+ D

Sbjct: 115 WGKFAAEIRDPAAKGSRIWLGTTFESDIDAARAYDYAAPKLRGRKAVLNPL--DAGKYDA 172

Query: 215 PVRISKR 222
PV KR
Sbjct: 173 PVNSCRKR 180

>G1791 (Amino Acid Sequence) (gf=1)
Length = 139

Score = 110 bits (274), Expect = 8e-26
Identities = 58/105 (55%), Positives = 66/105 (62%), Gaps = 10/105 (9%)

Query: 148 YGVGRQPGWKFAAEIRDPKNGARGWLGTFTETAEDAAALAYDRAAFPMRGSRALLNPFPLR 207
YGVGR+RPWGK+AAEIRD A++GARVWLGT? TAEDAA AYDRAAF MRG RA+LNFP
Sbjct: 14 YGVGRKPGWKYAAEIRDSARHGARGWLGTFTNTAEDAAARAAYDRAAFGMRCQRAILNPFPH 73

Query: 208 VNSGEPDPVRIKSKRSSFSNNENGAPKRRTVVAGGGMDKGLTV 252
+ P + S+EN GGG D G V

SDJCT: 74 YQMKDP-----NGSHENAVSSSSGYRGGGGDDGREV 108

>G1792 (Amino Acid Sequence) (gf=1)
Length = 139

Score = 107 Bits (268), Expect = 4e-25
Identities = 55/92 (59%), Positives = 69/92 (74%), Gaps = 5/92 (5%)

Query: 144 KGNHTRGVQRPRGWKFAAEIRDPFANGAKARVWLGTFETAGHAAALADRAAPNRGSGKALLN 200
K +RGVR+RPWGKFAAEIRDP++NGAR+WLGTPTETAA+AA AYDRAAP +RG A+LN
Sbjct: 16 KQARFGVRRRPWGKFAAEIRDPSPNGARLWLGTPTETAAARAYDRAAPNLRGHLAILN 75

Query: 204 PPLRVNNGEP--DPVRIKSKRSSPSSSNNENGA 233
 PP N P D ++ +S SSS+ +G+
 Spict: 76 PP---NEYVPRMDYSLRPPYASSSSSSSSSSGS 104

```
>Cl841 (Amino Acid Sequence) (gf=1)
```

Length = 229
Score = 107 bits (267), Expect = 5e-25

Query: 94 GWEPSSSSDEDRSFPVKIETPESFAAVDSV--PVKKEKTPVSAAVTAAGKHVRGV 151
Identities = 66/139 (47%), Positives = 80/139 (57%), Gaps = 4/139 (2%)

G SSSS+ R P SFA VDS P++ E S K + YRGV
 Sbjct: 31 GAHSSSSAGHKRGWLGIDSAPISPFARVDSSHPIE-ESMSKAPPEAREKKRRYRGV 89

Query: 152 RQRPWGAFAAETKDFPANGARVWILGTFETAEDAAAYDKAAFRNKGSKALNFPKRVNSG 211
RQRPWGK+AAEIRDP + ARVWILGTF+TAE AA AYD AA R RG++A LNPP V
Sbjct: 90 RQRPWGKAAEIRDPHR-AARVWILGTFDTAEAAARAYDEAALRPGNKAKLNPPEDVRIL 148

Query: 212 EDPVRIKSRSPSSSNE 230
 P P ++S + ++ E
 Spct: 149 PPPPLLRSPADTVAKAE 167

>G2512 (Amino Acid Sequence) (qf=1)

```
Length = 244
Score = 107 bits (266), Expect = 6e-25
```

Identities = 64/143 (44%), Positives = 82/143 (56%), Caps = 16/143 (11%)
Query: 95 WEPSSSSDEDRSSPPSVKIETP-ESFAADVSPVKKEKTPSVSAANTAAKGKHGVRQ 153

WE S D-S S + + F A + S +K+E AA + K YRGVR+
 Sbjct: 30 WEESPLHQSFQDQSLSSPTDNYCDDPPAPESSIIKEEGKATVAA--BEEKSYRGVRK 87

Query: 154 RPKGFAAIEIRDPAKNGARVWLGTFTETAEDAAALAYDRAAFMRGSRALLNFP----- 205
RPKGFAAIEIRD + G RVMLGTF+TAE AALAYD+AAF ++GS A+LNFP
Sbjct: 88 RPKGFAAIEIRDSTRKIGVRLGTFDTAEAAALAYDQAALFALGSLAVLNFPADVVEESL 147

Query: 206 -----LRVNSCEPDVPVRIKSRK 223
+ N CE + +K K S
Sbjct: 148 RKNENVLNDGESPVALRKHS 170

>G1419 (Amino Acid Sequence) (gf=1)
Length = 221

Score = 107 bits (266), Expect = 6e-25
Identities = 50/91 (54%), Positives = 66/91 (71%), Gaps = 5/91 (5%)

Query: 146 KHYGVQRQPMGKFAAIEIRDPKNGARVWLGTFTETAEDAAALAYDRAAFMRGSRALLNFP 205
+HYGVQR+RPMGK+AAIEIRDP K G RVMLGTF+TA +AA YD+AAF+RGS+A+LNFP
Sbjct: 70 RHYGVRRRPMGKYAAIEIRDPNKKGVRLGTFDTMEAGYDCKAAYKLGRSKAILNFP 129

Query: 206 LRVNSGE-----PDVRIKSRKSSFSNNEN 231
L E + +K+KR + +B+
Sbjct: 130 LEAGKHEDLGDNKKTISLKAKEKQVTEDES 160

>G1411 (Amino Acid Sequence) (gf=1)
Length = 248

Score = 104 bits (260), Expect = 3e-24
Identities = 56/113 (49%), Positives = 73/113 (64%), Gaps = 10/113 (8%)

Query: 118 ESPAAVDSVPVKKETSPVSAAVTAAGKHGVRQPMGKFAAIEIRDPKNGARVWLGT 177
+++ +V + ++ +P + +HYGVQRQPMGK+AAIEIRDP K ARVWLGT
Sbjct: 60 DNPVSVPDQPNQVAPTHQDQDRLRRHYGVQRQPMGKAAIEIRDP-KKAARVWLGT 118

Query: 178 FETAEDAAALAYDRAAFMRGSRALLNPLRVNSGEPDVPVRIKSRKSSFSNNEN 230
FETA E AALAYD AA + +GS+A LNFP RV G S++ SSN+
Sbjct: 119 FETAESAAALAYDEAALKPKGSKALNFPERVOLG-----SNSTYSSNQ 162

>G865 (Amino Acid Sequence) (gf=1)
Length = 212

Score = 104 bits (260), Expect = 3e-24
Identities = 54/84 (64%), Positives = 64/84 (75%), Gaps = 3/84 (3%)

Query: 127 PVKKE--KTSVSAAVTAAGKHGVRQPMGKFAAIEIRDPKNGARVWLGTFTETAED 184
PVK+E K+ + +HYGVQRQPMGK+AAIEIRDP K ARVWLGTFTETA E A
Sbjct: 16 PVKBLDKSDQHOPDQDQPRRHRYGVQRQPMGKAAIEIRDP-KKAARVWLGTFTETA E A 74

Query: 185 ALAYDRAAFMRGSRALLNPLRV 208
ALAYDRAA + +G+ A LNFP RV
Sbjct: 75 ALAYDRAALFKGTAKALNFPERV 90

>G30 (Amino Acid Sequence) (gf=1)
Length = 133

Score = 103 bits (257), Expect = 7e-24
Identities = 50/70 (71%), Positives = 57/70 (81%), Gaps = 1/70 (1%)

Query: 142 AKGKHGVRQPMGKFAAIEIRDPKNGARVWLGTFTETAEDAAALAYDRAAFMRGSRAL 201
A +GK YGVQR+RPMGK+AAIEIRD K+G RVMLGTF+TAE DRAA AYDRAA+ MRG A+
Sbjct: 15 AEOGK-YGVRRRPMGKYAAIEIRDSRKHGERVWLGTFTETAEDAAALAYDRAAAYMRGKAAI 73

Query: 202 LNFPLRVNSG 211

LNFP N G
Sbjct: 74 LNFPPHYNMG 83

>G44 (Amino Acid Sequence) (gf=1)
Length = 241

Score = 103 bits (256), Expect = 9e-24
Identities = 54/113 (47%), Positives = 73/113 (63%), Gaps = 10/113 (8%)

Query: 146 KHYGVQRQPMGKFAAIEIRDPKNGARVWLGTFTETAEDAAALAYDRAAFMRGSRALLNFP 205
+HYGVQR+RPMGK+AAIEIRDP K G RVMLGTF+TA +A AYD+AAP++RG +A+LNFP
Sbjct: 86 RHYGVRRRPMGKYAAIEIRDPNKKGCRILGTYDTAVEAGRAYDQAALFALGSKAILNFP 145

Query: 206 LRVNSGEPDVPVRIKSRKSSFSNNENGAPKKRTVAAGGMDKGLTVKCEVE 258
L VR+ S+ + S G K+R+ + +K VK E E
Sbjct: 146 L-----DVRVTSE--TCSGEGVIGLKKRDKGSPPEEKAARVKVEEE 188

>G1795 (Amino Acid Sequence) (gf=1)
Length = 131

Score = 100 bits (250), Expect = 5e-23
Identities = 52/97 (53%), Positives = 67/97 (68%), Gaps = 1/97 (1%)

Query: 140 VTAAGKHGVRQPMGKFAAIEIRDPKNGARVWLGTFTETAEDAAALAYDRAAFMRGSR 199
V A GK YGVQR+RPMGK+AAIEIRD K+G RVMLGTF+TAE+AA AYD+AA+ MRG
Sbjct: 8 VGAEHCK-YGVRRRPMGKYAAIEIRDSRKHGERVWLGTFTETAEDAAALAYDQAALFALGSKAILNFP 66

Query: 200 ALLNPLRVNSGEPDVPVRIKSRKSSFSNNENGAPKK 236
A+LNFP N G SS +S++ + + + +

Sbjct: 67 AILNFPHEYNMGSGVSSSTAMAGSSASASASSSRQ 103

>G1385 (Amino Acid Sequence) (gf=1)
Length = 336

Score = 100 bits (249), Expect = 6e-23
Identities = 60/138 (43%), Positives = 78/138 (56%), Gaps = 10/138 (7%)

Query: 110 PSVKIETPESPAVDSVPVKKETSPVSAAVTAAGKHGVRQPMGKFAAIEIRDPKRN 169
PS ++ +P+ + P SA K YGVQRQK W GK+ AEIR P +N
Sbjct: 155 PSGRMMMLRQESDLPLTRPVPQPSAT-----KLYGVQRQKHWKVAEIRKP-RN 207

Query: 170 CARVWLGTFTETAEDAAALAYDRAAFMRGSRALLNFP-LRVNSGEPDVPVRIKSRKSSPS- 227
R+WLGTFTETA E A+AA+AYDR AP++RG A LNFP L +N EP PV K + +S
Sbjct: 208 RTRLWLGTFTETA E A+AA+AYDR AP++RG A LNFP L +N EP PV K + +S 267

Query: 228 -SNENGAPKKRTVAAGG 244
S+ G +A GG
Sbjct: 268 DSSRRGEDSDSTALAVGG 285

>G1752 (Amino Acid Sequence) (gf=1)
Length = 243

Score = 100 bits (248), Expect = 8e-23
Identities = 50/89 (56%), Positives = 63/89 (70%), Gaps = 7/89 (7%)

Query: 146 KHYGVQRQPMGKFAAIEIRDPKNGARVWLGTFTETAEDAAALAYDRAAFMRGSRALLNFP 205
+ YGVQR+RPMGK+AAIEIRD +NG RVMLGTF+ AE+ AALAYD+AAP +GS A LNFP
Sbjct: 84 RSYGVRRRPMGKFAAIEIRDSRNGIRVWLGTFDKABEAAALAYDQAALFALGSKAILNFP 143

Query: 206 LRVNSGEPDVPVRIKSRKSSFSNNENGAP 234
+ V VR K+ + ++ G+P

Query: 201 LLNPLRVNSGEPDPVR 217
L NP ++ EP+PV+
Sbjct: 170 LTNFTVQ---PEPEPVQ 183

>G864 (Amino Acid Sequence) (gf=1)
Length = 343

Score = 98.6 bits (244), Expect = 2e-22
Identities = 58/135 (42%), Positives = 79/135 (57%), Gaps = 19/135 (14%)

Query: 128 VKKEKT-SVSAVAATAAG-KHYRGVQRPWGKFAAEIRDPKNGARVWLGTFTETADAA 185
VK E T SPV +A T G K +RGVQRPWGK+AAEIRDP K R+WLGT+ TAE+AA
Sbjct: 100 VKSESTSVSVSATTTTTEKKKFGVQRPWGKAAEIRDELKR-VRLWLGTNTAEAA 158

Query: 186 LAYDRAAPRMGRSALLNPPPLRVNSG-----EPDPVRIKSKSSSFSSSNENGAPEKRRRT 239
+ VD AA +RG AL NP + + P PV+ K K+++ K +++
Sbjct: 159 MYVDNAALQLRCPDALTNSVPTTATEKKAPPPSPVKKKKKNN-----KSKKS 208

Query: 240 VAAGGGMKGLTVKC 254
V A + + + C
Sbjct: 209 VTASSISRSSSND 223

>G4 (Amino Acid Sequence) (gf=1)
Length = 375

Score = 97.8 bits (242), Expect = 4e-22
Identities = 63/150 (42%), Positives = 84/150 (56%), Gaps = 17/150 (11%)

Query: 91 PHGWPSSSSSDSDRSFSPVK-----IETPSPAIVDSVPVKKEKTSFVSAVTAAG 144
P G + S+ ++D F +VK P + A V +V K + SA
Sbjct: 56 FQGFDDSAFDCEDDDVFVNVKPFVFTATTKPVASAFVSTVGSAAKTVESAEQAKS 115

Query: 145 [KH]-----YRGVQRPWGKFAAEIRDPKNGARVWLGTFTETADAAALAYDRAAFMRGSR 199
K YRG+RORPWGK+AAEIRDP K G+R WLGTFTAE+AA AYD AA R+RG++
Sbjct: 116 SKRKKNQYRGVQRPWGKAAEIRDPK-GSREWLGTFTAEAAAYDAAARRIRGTK 174

Query: 200 ALLNPLRVNSGEPDPVRIRKSSSFSSSN 229
A +NFP E +P + KR S ++N
Sbjct: 175 AKVNFP-----BEKNPSPVQKPSAKTNN 199

>G2130 (Amino Acid Sequence) (gf=1)
Length = 315

Score = 96.7 bits (239), Expect = 9e-22
Identities = 57/119 (47%), Positives = 71/119 (58%), Gaps = 14/119 (11%)

Query: 99 SSSSDSDRSF---PSVKI-----ETPSPAIVDSVPVKKEKTSFVSAVTAAG 145
SSS D+D ++ P VK E+ S AA ++E T+ K
Sbjct: 43 SSSDDDDNNVTVPVRKYVKEIRPCOGSSSTAAARKKHKEEESVVEDDSTSVKP 102

Query: 146 KHYRGVQRPWGKFAAEIRDPKNGARVWLGTFTETADAAALAYDRAAFMRGSRALLNF 204
K YRGVQRPWGKFAAEIRDP+ R+WLGTFTAE+AA+AYDRAA ++G +AL NF
Sbjct: 103 KKYRGVQRPWGKFAAEIRDPSSR-TRILGTFTVTAEEAAIAYDRAAILHKGKALTNF 160

>G7 (Amino Acid Sequence) (gf=1)
Length = 192

Score = 96.7 bits (239), Expect = 9e-22
Identities = 49/78 (62%), Positives = 57/78 (72%), Gaps = 1/78 (1%)

Sbjct: 144 VEV-----VRESLKQMNVLHDGSP 165

>G1751 (Amino Acid Sequence) (gf=1)
Length = 268

Score = 100 bits (248), Expect = 8e-23
Identities = 62/125 (49%), Positives = 72/125 (57%), Gaps = 5/125 (4%)

Query: 124 DSVPVKKEKTSFVSAVAATAAGKHYRGVQRPWGKFAAEIRDPKNGARVWLGTFTETAD 183
+S PV K+ K YRGVQRPWGKFAAEIRDP K RVWLGTFTETAD
Sbjct: 112 ESSPVAKAAGGGKIRKKNKNGYRGVQRPWGKFAAEIRDP-KRATRVWLGTFTETAD 170

Query: 184 ALAYDRAAPRMGRSALLNPP-LRVNSGEPDPVR---TKSRKSSSFSSSNENGAPEKRRRT 239
AA AYDRAA RG RA LNPP + S PV I +K S+ +S + + +
Sbjct: 171 AARAYDRAALGFRGKRAKLNPPFDVTSVSSPVAADDIGAKASASASVSAITDSVEAQW 230

Query: 240 VAAGG 244
GG
Sbjct: 231 NGGGG 235

>G2509 (Amino Acid Sequence) (gf=1)
Length = 263

Score = 99.4 bits (246), Expect = 1e-22
Identities = 48/62 (77%), Positives = 53/62 (85%), Gaps = 1/62 (1%)

Query: 146 KHYRGVQRPWGKFAAEIRDPKNGARVWLGTFTETADAAALAYDRAAFMRGSRALLNPP 205
+HYRGVQRPWGK+AAEIRDP K ARVWLGTFTAE AALAYD AA + +GS+A LNPP
Sbjct: 90 RHYRGVQRPWGKAAEIRDPK-AARVWLGTFTETAEAAALAYDRAALKFKGSKAKLNPP 148

Query: 206 LR 207
R
Sbjct: 149 ER 150

>G26 (Amino Acid Sequence) (gf=1)
Length = 218

Score = 99.4 bits (246), Expect = 1e-22
Identities = 49/69 (71%), Positives = 59/69 (85%), Gaps = 3/69 (4%)

Query: 143 AKGKHYRGVQRPWGKFAAEIRDPKNGARVWLGTFTETADAAALAYDRAAFMRGSRALL 202
++ ++YRGVQRPWGK+AAEIRDP K ARVWLGTFTAE+AALAYD+AAF RG +A L
Sbjct: 65 SRQRNRYGVRQRPWGKAAEIRDPK-AARVWLGTFTETAEAAALAYDKAAFEPRGHKAKL 123

Query: 203 NFP--LRVN 209
NFP ++RVN
Sbjct: 124 NPFHIRVN 132

>G1750 (Amino Acid Sequence) (gf=1)
Length = 335

Score = 98.6 bits (244), Expect = 2e-22
Identities = 58/137 (42%), Positives = 80/137 (58%), Gaps = 19/137 (13%)

Query: 96 EPPSSSDSDRSFSPVKIE-----ETPSPAIVDSVPVKKEKTSFVSA-----AV 140
+ + SSSDE+ FP +++ + S V + P K K V + +V
Sbjct: 51 DATSSDDEEFLLFPKRRKLINEIRVPEPSSSTGDVSPKTKRKIRINVDSTVQKPSV 110

Query: 141 TAAKGKHYRGVQRPWGKFAAEIRDPKNGARVWLGTFTETADAAALAYDRAAFMRGSR 200
+ K YRGVQRPWGK+AAEIRDP + R+WLGTFTAE+AA+ YD AA ++RG A
Sbjct: 111 SGONOKYRGVQRPWGKAAEIRDPQR-RRILGTFTAEAAIAYDRAAILHKGKALTNF 169

43/19 = 6231

Query: 133 TSPVSAVTAAGKHGVRGVRQPMGKFAAEIRDPAKNGARVWLGTFTAEADALAYDRAA 192
T + + + + K YRGVRQPMGK-AAEIRD P K RVWLGTFTAE AA AYD AA
Sbjct: 46 TXCESSNLERPKYRGVRQPMGKFAAEIRDPAKNGARVWLGTFTAEADALAYDRAA 104

Query: 193 PMRGSRALLNFPRLVNS 210
Sbjct: 105 LRFPGSKAKLNFPENVGT 122

>G45 (Amino Acid Sequence) (gf=1)
Length = 349

Score = 96.3 bits (238), Expect = 1e-21
Identities = 56/133 (42%), Positives = 73/133 (54%), Gaps = 27/133 (20%)
Query: 107 SFPSPVKIETPESFAAVDSVPVKKEKTSFVSAVTAAGKHGVRGVRQPMGKFAAEIRD P 166
+SFP+ P + +DS + +K + YRGVR+PMGKFAAEIRD
Sbjct: 124 ASPFNKTHDPLNPTIDSCLSR-----VSKTKRYGVRKPMGKFAAEIRDS 174
Query: 167 AKNGARVWLGTFTAEADALAYDRAAPMRGS-RALLNFPRLV----- 208
+NG RVWLGTFTAE+AA+AYD+AA R+RG+ +A NF L
Sbjct: 175 TRNGVRWLGTFTAEADALAYDRAAVIRGTQKHTNFQLETVIKAMEMDCNPYRNN 234

Query: 209 NSGEPDPVRIKSK 221
NS DP+R K
Sbjct: 235 NSNTSDPLRSRK 247

>G19 (Amino Acid Sequence) (gf=1)
Length = 248

Score = 96.3 bits (238), Expect = 1e-21
Identities = 60/141 (42%), Positives = 79/141 (55%), Gaps = 6/141 (4%)
Query: 95 WEFSSSDESSSPSPV-----RSSPFSVKIETPE---SFAAVDSVPVKKEKTSFVSAVTAAGKHGVROR 154
W +S+ +D F S P + V VKKE+ + YRG+R+R
Sbjct: 28 WSELDASADDFPMGFYSTSKLHPTNQVNVKEAVKKEQAT--EPGKRKKNYRGIRKR 85
Query: 155 PMGKFAAEIRDPAKNGARVWLGTFTAEADALAYDRAAPMRGS-RALLNFPRLVNSGEH 214
PMGK+AAEIRD P K G RVWLGTFTAE+AA+AYD AA +RG +A LNFP + P
Sbjct: 86 PMGKFAAEIRDPAK-GVRWLGTFTAEADALAYDRAAPMRGS-RALLNFPRLVNSGEH 214

Query: 215 PVRIKSKSSSSSSSNGAPK 235
P SS S+++ A K
Sbjct: 142 PPNYTPPSPSPRSTQPPAKK 162

>G14 (Amino Acid Sequence) (gf=1)
Length = 358

Score = 95.5 bits (236), Expect = 2e-21
Identities = 62/127 (48%), Positives = 77/127 (59%), Gaps = 13/127 (10%)
Query: 91 FHGMEPSSSSSDED-----RSSPFSVKIETPE---SFAAVDSVPVKKEKTSFVSAAV 140
P G + SS D+D P V TP+ S AA SV KK A
Sbjct: 56 FQGFKDDSSIDCDDDFADVADVPVFTSTPKPAVSAABSGSVFGKVTGLDGAERK 115
Query: 141 TAAKCK--HYRGVRQPMGKFAAEIRDPAKNGARVWLGTFTAEADALAYDRAAPMRGS 198
+A + + YRG+RORPMGK+AAEIRD P GAR+WLGTFTAE+AA AYD AA R+RGS
Sbjct: 116 SANRKRKQYRGIRQPMGKFAAEIRD-REGARWLGTFTAEADALAYDRAAPMRGS 174

Query: 199 RALLNFP 205
+A +NFP
Sbjct: 175 KAKWNP 181

>G1008 (Amino Acid Sequence) (gf=1)
Length = 294
Score = 94.4 bits (233), Expect = 4e-21
Identities = 55/117 (47%), Positives = 66/117 (56%), Gaps = 9/117 (7%)
Query: 96 EPSSSSSDESSSPSPVKIETPE-SFAAVDSVPVKKEKTSFVSAVTAAGKH----- 147
+ SS +S+ +S P VK E F D P K S A A G
Sbjct: 39 DSSSDEENDKNSVAPRVKRYVDIEIRPCDEDEPKAPAKKKSPAAAEENGCDLVKSVVK 98
Query: 148 YRGVRQPMGKFAAEIRDPAKNGARVWLGTFTAEADALAYDRAAPMRGS-RALLNFP 204
YRGVRQPMGKFAAEIRD P +R+WLGTFTAE+AA+ YDRAA R+RG A NF
Sbjct: 99 YRGVRQPMGKFAAEIRDPSR-TRLWLGTFTAEADALAYDRAAPMRGS-RALLNFP 154

>G1421 (Amino Acid Sequence) (gf=1)
Length = 287

Score = 94.0 bits (232), Expect = 6e-21
Identities = 56/121 (46%), Positives = 76/121 (62%), Gaps = 16/121 (13%)
Query: 99 SSSSDESSSPSPV-----KIETPESPAVDSV-----PVKKEKTS---PVSAAVTAAK 144
SSS +E+ F + + + +DSV P+KK+ V VT A
Sbjct: 25 SSSDEEEVDFALSTKRERVKYKVEVLDSVSDKPKMKKKRKRVRVTVVVVITAT 84
Query: 145 GKHYRGVRQPMGKFAAEIRDPAKNGARVWLGTFTAEADALAYDRAAPMRGS-RALLNFP 204
K +RGVRQPMGK+AAEIRD P +RVWLGTFTAE+AA+ YD AA +RG A LNFP
Sbjct: 85 RK-FRGVRQPMGKFAAEIRDPSR-VRWLGTFTAEADALAYDRAAPMRGS-RALLNFP 142
Query: 205 P 205
P
Sbjct: 143 P 143

>G1794 (Amino Acid Sequence) (gf=1)
Length = 391

Score = 93.6 bits (231), Expect = 7e-21
Identities = 51/89 (57%), Positives = 59/89 (65%), Gaps = 6/89 (6%)
Query: 146 KHYRGVRQPMGKFAAEIRDPAKNGARVWLGTFTAEADALAYDRAAPMRGS-RALLNFP 205
+ YRGVRQPMGK+AAEIRD P K ARWLGTFTAE AA AYD AA R RG+A LNFP
Sbjct: 183 RYRGVRQPMGKFAAEIRDPAKNGARVWLGTFTAEADALAYDRAAPMRGS-RALLNFP 241
Query: 206 -----LRVNSGEPDPVRIKSKSSSPSSN 229
+R S E PV + + S N
Sbjct: 242 ENVKLYRPASTEAPVHQTAAQRPTQSRN 270

>G440 (Amino Acid Sequence) (gf=1)
Length = 354

Score = 92.8 bits (229), Expect = 1e-20
Identities = 66/177 (37%), Positives = 88/177 (49%), Gaps = 31/177 (17%)
Query: 56 SPSKLYPCFTESGDLPLKNDSEDML-----VYGLINDAPHGWEPPSS-----SSSD 103
S K+ C+T+ + D ED L V +N+ EPS S D
Sbjct: 35 SLKVCITDIDATDSSDEDEDFLPFRVRKRVNEITV---EPSCNNVTVGVSKMD 91
Query: 104 EDRSSPSPVKIETPESPAVDSVPVKKEKTSFVSAVTAAGKHGVRQPMGKFAAEI 163
R S S + +P S + P + + K +RGVRQPMGK+AAEI
Sbjct: 92 RKLSSSSDETQSPAS-----SRQPNKNSVSGQKKKFRGVRQPMGKFAAEI 140
Query: 164 RDPKNGARVWLGTFTAEADALAYDRAAPMRGS-RALLNFP-----RVNSGEPDPV 216

Score = 89.4 bits (220), Expect = 1e-19
 Identities = 56/122 (45%), Positives = 70/122 (56%), Gaps = 7/122 (5%)
 Query: 133 TSPVSAVTAAGKHGVRQRPWGKFAAEIRDPAKNGARVWLGTFTAEADAALAYDRAA 192
 TS V K YRGVROR WGR+ AEIR P KN R+WLGTFTAE+AAALAYD+AA
 Sbjct: 98 TSMKKIDVATKPVKLYRGVRQRPWGKFAAEIRLP-KNRTLMGLTFTAEQAALAYDQAA 156
 Query: 193 FMRGSRALLNPLRVNSGEPDVP---RKSRSFSSSSNENGAP---KRETVTAAGSGM 246
 ++RG A LNPP V C + I +K S +S++ P K +T C
 Sbjct: 157 HKRGDARLNFPDIVRGHYKQILSPSINAKIESICNSDLPOLPEKQNKTEEVLSGF 216
 Query: 247 DK 248
 K
 Sbjct: 217 SK 218

>G2135 (Amino Acid Sequence) (gf=1)
 Length = 262
 Score = 88.2 bits (217), Expect = 3e-19
 Identities = 50/103 (48%), Positives = 65/103 (62%), Gaps = 13/103 (12%)
 Query: 103 DEDRSSPESVKIETPESFAAVDSVPVKKEKTSFVSAAVTAAGKHGVRQRPWGKFAAE 162
 DE S P V T + A + KK++S Y+G+R+RPMG++AAE
 Sbjct: 43 DELESDEPFPVFSSTHKHAGSGSDGKKQSS-----RYKGIIRRPWGRMAAE 90
 Query: 163 IRDPKAGCARVWLGTFTAEADAALAYDRAAPMRGSRALLNPP 205
 IRDP K G RVWLGTFTAE+AA AYD A R+RC+A LNPP
 Sbjct: 91 IRDPK-GVRWLGTFTAEAMANDLEAKKIRGAKAKLNPP 132

>G1380 (Amino Acid Sequence) (gf=1)
 Length = 244
 Score = 88.2 bits (217), Expect = 3e-19
 Identities = 45/79 (56%), Positives = 56/79 (69%), Gaps = 2/79 (2%)
 Query: 135 PVSAAVTAAGKHGVRQRPWGKFAAEIRDPAKNGARVWLGTFTAEADAALAYDRAAPR 194
 PV+A + K YRGVR+RPMG+FAAEIRDP K +RVWLGTFTAE+AA AYD AA
 Sbjct: 15 PVTAG-GSVKPYRGVRQRPWGKFAAEIRDPK-K-SRVWLGTFTAEADAALAYDRAAPR 72
 Query: 195 MRGSRALLNPLRVNSGEP 213
 +RG +A NPP+ + P
 Sbjct: 73 LRCPKAKTNFPIDCSFSP 91

>G625 (Amino Acid Sequence) (gf=1)
 Length = 328
 Score = 88.2 bits (217), Expect = 3e-19
 Identities = 63/130 (48%), Positives = 71/130 (54%), Gaps = 16/130 (12%)
 Query: 134 SPVSAVTAAGK-----HYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFTAEADA 185
 S +A KKG YRGVROR WGR+ AEIR+P K R WLGTFTAEADA
 Sbjct: 33 SSTSAQRKKGKGGGPDNSKPRYRGVRQRPWGKFAAEIRPEKR-TRKWLGTFTAEADA 91
 Query: 186 LAYDRAAFMRGSRALLNPF----PLRVNSGEPDVPRIKRSFSSSSNENGAPKKRTVA 241
 AYDRAA + GSRA LN P V+S V S S+ SSS + P R A
 Sbjct: 92 RAYDRAAVLYGSRALNTSSPSSVSSSS-SVSAASSPSTSSSTQTLRLPLPRAA 150
 Query: 242 A--GGGWDKG 249
 A GGG + G
 Sbjct: 151 ATVGGGANFG 160

>G1748 (Amino Acid Sequence) (gf=1)
 Length = 261
 Score = 87.4 bits (215), Expect = 5e-19
 Identities = 47/82 (57%), Positives = 57/82 (69%), Gaps = 2/82 (2%)
 Query: 148 YRGVRQRPWGKFAAEIRDPAKNGARVWLGTFTAEADAALAYDRAAPMRGSRALLNPP 207
 YRGVROR WGR+ AEIR P KN R+WLGTFTAE+AAALAYD+AA+RG A LNPP
 Sbjct: 71 YRGVRQRPWGKFAAEIRLP-KNRTLMGLTFTAEKAALAYDQAAFLRGDIKLNPPNL 129
 Query: 208 VNSG-BPDVPRIKRSFSSSS 228
 ++ P + +K + S
 Sbjct: 130 IHEDMNPSPSSVDTKLOAICKS 151

>G1845 (Amino Acid Sequence) (gf=1)
 Length = 292
 Score = 87.0 bits (214), Expect = 7e-19
 Identities = 41/60 (68%), Positives = 50/60 (83%), Gaps = 1/60 (1%)
 Query: 146 KHYRGVRQRPWGKFAAEIRDPAKNGARVWLGTFTAEADAALAYDRAAPMRGSRALLNPP 205
 K YRGVROR WGR+ AEIR P ++ AR+WLGTFTAE+AA+AYDR AF+RG A LNPP
 Sbjct: 141 KLYRGVRQRPWGKFAAEIRKP-RSARLWLGTFTAEAAAMAYDRAQFLRGHSATLNPP 199
 >G2059 (Amino Acid Sequence) (gf=1)
 Length = 343
 Score = 86.3 bits (212), Expect = 1e-18
 Identities = 54/111 (48%), Positives = 66/111 (58%), Gaps = 9/111 (8%)

Query: 130 KETSFPVSAVTAAGKHGVRQRPWGKFAAEIRDPAKNGARVWLGTFTAEADAALAYD 189
 K KT V T K YRGVROR WGR+ AEIR P +N RVWLGTFTAE+AA+AYD
 Sbjct: 171 KSKTRKVVQTTPTK--LYRGVRQRPWGKFAAEIRLP-RNRTVWLGTFTAEQAAMAYD 227
 Query: 190 RAAPMRGSRALLNPLRVNSGEPDVP-----IKSRSSFSSSSNENGAP 234
 AA+ +RG A LNPP + +R ++SK SSS + +P
 Sbjct: 228 TAAYILRGEPALNFPDLKHLKSGSLRCMIASLLESKIQIISSSQVSNP 278

>G1020 (Amino Acid Sequence) (gf=1)
 Length = 185
 Score = 86.3 bits (212), Expect = 1e-18
 Identities = 45/78 (57%), Positives = 53/78 (67%), Gaps = 1/78 (1%)
 Query: 133 TSPVSAVTAAGKHGVRQRPWGKFAAEIRDPAKNGARVWLGTFTAEADAALAYDRAA 192
 T+P AKK YRGVR+RPMG+AAAEIRDP K RVWLGTFTAE+AA AYD AA
 Sbjct: 16 TNPTHESNAKAEIRYRGVRQRPWGKFAAEIRDPVK-TRVWLGTFTAEQAAMAYDAA 74
 Query: 193 FMRGSRALLNPLRVNS 210
 RG +A NF + V S
 Sbjct: 75 RDPGRVAKTNFVGIVGS 92

>G6 (Amino Acid Sequence) (gf=1)
 Length = 222
 Score = 86.3 bits (212), Expect = 1e-18
 Identities = 42/63 (66%), Positives = 49/63 (77%), Gaps = 1/63 (1%)
 Query: 143 AKGKHGVRQRPWGKFAAEIRDPAKNGARVWLGTFTAEADAALAYDRAAPMRGSRALL 202
 AK YRGVR+RPMG+AAAEIRDP K RVWLGTFTAE+AA AYD AA RG++A

Sbjct: 20 AKEIRYGVRRPWRGKFAAEIRDPGKK-TRVWLGTFTDAEEAARAYDTAARDPRGAKAKT 78
Query: 203 NPP 205
NPP
Sbjct: 79 NPP 81

>G1796 (Amino Acid Sequence) (gf=1)
Length = 328
Score = 85.9 bits (211), Expect = 2e-18
Identities = 52/128 (40%), Positives = 66/128 (50%), Gaps = 9/128 (7%)
Query: 110 PSVKLETPEGFAAVDSVPVKEKTSVSAAVTAAGKHGVRQRPWGKFAAEIRDPAKN 169
P K T + + V K + V + A + YRGVR+RPMG+AAEIRDP
Sbjct: 19 PLTKFTPTASPVSNRKLSSKDTVTIAGAGSITTRYGVRRRPWRGKFAAEIRDPMSK 78

Query: 170 GARVWLGTFTDAEDALAYDAAPRMRGSRALLNPLRPNVRIKSKSPSSSN 229
R WLGTFTAE AA AID AA RG+AA NF P+P PS SN
Sbjct: 79 ERR-WLGTFTDAEQAAACAYDAANAFRGAKNTFTTYTAVIMPEP-----RFSFSN 129

Query: 230 ENGAPKKR 237
+ +P R
Sbjct: 130 KKSSPSAR 137

>G5 (Amino Acid Sequence) (gf=1)
Length = 334
Score = 85.9 bits (211), Expect = 2e-18
Identities = 42/68 (61%), Positives = 53/68 (77%), Gaps = 1/68 (1%)
Query: 138 AAVTAAGKHGVRQRPWRGKFAAEIRDPKNGARVWLGTFTDAEDALAYDAAPRMRG 197
+ V + K YRGVROR WGR+ AEIR P +N R+WLGTFT+TAE+AAAYD-AA+++RG
Sbjct: 142 SGVPSKTKLYRGVRORHWGKVAEIRLP-RNRTRLWLGTFTDAEDALAYDAKAAAYKLRG 200

Query: 198 SRALLNPP 205
A LNPP
Sbjct: 201 DFARLNPP 208

>G2113 (Amino Acid Sequence) (gf=1)
Length = 166
Score = 85.1 bits (209), Expect = 3e-18
Identities = 39/61 (63%), Positives = 48/61 (77%), Gaps = 1/61 (1%)
Query: 145 GKHYGVRQRPWKFAAEIRDPKNGARVWLGTFTDAEDALAYDAAPRMRGSRALLNPP 204
+ V + K YRGVROR WGR+AAEIRDP K +RVWLGTFT+T E+AA AYD+ A RG+AA NF
Sbjct: 17 GVRVGRKRPWGRYAAEIRDPK-KSRVWLGTFTDTEEAARAYDKRAIEFRGAKAKTNP 75

Query: 205 P 205
P
Sbjct: 76 P 76

>G2571 (Amino Acid Sequence) (gf=1)
Length = 335
Score = 84.7 bits (208), Expect = 3e-18
Identities = 43/70 (61%), Positives = 52/70 (73%), Gaps = 1/70 (1%)
Query: 136 VSAVTAAGKHGVRQRPWRGKFAAEIRDPKNGARVWLGTFTDAEDALAYDAAPRMRG 195
+ +A K K YRGVROR WGR+ AEIR P +N RVWLGT+TAE AA AYDRAA+++
Sbjct: 124 LAAGRVTKKKLYRGVRORHWGKVAEIRLP-QNRMRVWLGTDTDAEAAAYAYDRAAYKL 182

G28 (Length: 268)

>G23 (Amino Acid Sequence) . (gf=1)
Length = 236

Score = 80.1 bits (196), Expect = 8e-17
 Identities = 44/109 (40%), Positives = 68/109 (62%), Gaps = 1/109 (0%)

Query: 97 PSSSSSDRRSSPPSVKVIETPESFAAVDSVPVKKETSIPSAVTAAKGKHVGVYORPM 156
 P+SSSSDQ++ +E++ +K+ S++ + + YRGV R W
 Sbjct: 10 P+SSSSDQATTTTTHLSEAPPPNNRRRSDSSASSSSQHPVYGVYGRMSM 69

Query: 157 GKPAEEDIPAKGAKVWLGTETPEAALAYDRAAPRMGRSALNFP 205
 GK+ EIR P K R WLGTF TA+ AA A+D AA +G S A+LNF
 Sbjct: 70 GKWSVEIROPKK- TRIGLITFTVTDMAARDAALATIKGSSAVLNFP 117

```
>G12 (Amino Acid Sequence) (gf=1)
      Length = 196
```

Score = 80.1 bits (196), Expect = 8e-17
Identities = 43/92 (46%), Positives = 55/92 (59%), Gaps = 6/92 (6%)

Query: 114 IETPESAAVDSPVKKEKTSVSAATAAKGKHVGRQRPWGKPAAEIRDPKNGARV 173
+ET A V S P V+A T + K Y+G+R R WKG+ ABIR+P K +R+
Sbjct: 1 METATEVATVVSPTAVT-----VAAVATRKRPYKGIKRWKKGWVAEIRSPNKR+SRI 54

Query: 174 WLGPETAEDAAALAYDRAAPMRGSRALLNFP 205
 WLG+ T E AA AYD A F +RG A LNFP
 Sbjct: 55 WLGSYSTPEAAARAYDTAVFYLRGSPARLNFP 86

```
>G1141 (Amino Acid Sequence) (gf=1)
      Length = 330
```

Score = 79.7 bits (195), Expect = 1e-16
Identities = 42/63 (66%), Positives = 46/63 (72%), Gaps = 1/63 (1%)

Query: 148 YRGVQRPWGKFAAEIRDPAKNGARVWLGTFTAEDAAALAYDRAAPRMGRSALLNPPLR 207
+RGVQR W GK+ AEIR+P K G R+WLGT TAE AA AYD AA M GS A LNPP
Sbjct: 78 FRGVQRINGKVAEIREP-KIGTRLWLGTFTFAEKAASAYDEAATAMVGSRLARLNPPQS 136

Query: 208 VNS 210
V S
sbjct: 137 VGS 139

>G2138 (Amino Acid Sequence) (gf=1)
Length = 161

Score = 79.3 bits (194), Expect = 1e-16
 Identities = 54/141 (38%), Positives = 68/141 (47%), Gaps = 22/141 (15%)

Query: 96 EPSSSDSEDRSS-----FPSVKIETPSFAANDSPVKEKTPSVSAVT 141
 + I + S + D + + P T
 E + SSSDEDTEERGASQTRRRGKRLVKEIVIDPSDSADKLDVCKTRFKIRPAFLKT 72

Query: 142 AAKGKHVGRVORPWGKPAEIRDPAKNGA-----RVWLGTFTETAEADAAALAYDRAAFRM 195
A K YRGVORPWGK+ AEIR GA R+WLGT TAE+AAALAYD A+ ++
Sbjct: 73 AKTEKHYGRVORPWGKQWAEIR--CGGACKRRDRDLWLGTFTETAEAAALAYDNASIKL 130

Query: 196 RGSRALNPLRVNSGEPDV 216
G A N F L + E V
Sbjct: 131 IGPHAPTNGFLPAENQEDTV 151

```
>G38 (Amino Acid Sequence) (gf=1)
      Length = 335
```

```

Score = 79.3 bits (194), Expect = 1e-16
Identities = 45/91 (49%), Positives = 55/91 (59%), Gaps = 9/91 (9%)

Query: 123 VDSVPVKKETKSPVSAVTAAGK-----HYKVRQRPMPKFAAEJRDPAKNGARV 174
      V + V KK K + GK +RGVRQR WGR+ AEIR+P + G+R+W
Sbjct: 46 VEEVTEKKRKVPAGSKCKMGCKGGPENSRCRSFVRQIRWGKVAIEEPR- GSRLL 104

Query: 175 LGTFPTAEADALVDRAAFAFRMGSRALLNFP 205
      LGTF TA+AA AYD AA M G A LNPP
Sbjct: 105 LGTFPTAOEAAASAVDEAAKAMYGPLARLNFP 135

```

>G975 (Amino Acid Sequence) (gf=1)
Length = 199

Score = 79.0 bits (193), Expect = 2e-16
 identities = 45/89 (50%), Positives = 54/89 (60%), Gaps = 2/89 (2%)

Query: 146 KHYRGVORPCKGFAEIRIDPAKNGAIVLWGHFTETAEADALAVDRAAPMGSGALLNFP 205
 K + KRGVGR WG + AEIR P + WIGHFTETAE + AA AYD RA M G A NFP
 Sbjct: 5 KXRGVGRGHWGSAVEIRHPILK - RRIWLGFTETAEAAVDEAAVLMSGRNAKTNFP 63

Query: 206 L -RVNSGEGDPVRIKKSRTFSFSSNENGA 233
 L N+GE SS SS+ + +

Sbjct: 64 LANNNTGTSBKTDISASSTWMSSTSS 92

>G1800 (Amino Acid Sequence) (gf=1)
Length = 277

Score = 78.2 bits (191), Expect = 3e-16
 Identities = 37/58 (63%), Positives = 45/58 (76%), Gaps = 1/58 (1%)
 Query: 148 YRGVQRQWCKPAAEIRDPKNGARVWLCTTETDADALYDRAAFPMRGSRALNFP 205
 YRGVQR WCK+ AEIR+P K AR+LWLS+G F TAE+AA+AYD AA ++ G A LN P
 Sbjct: 28 YRGVQR QWCKVWAEIRPKPKR AR+LWLSGATAEAAAYDAAALYGDHAYLNP 84

>G442 (Amino Acid Sequence) (gf=1)
Length = 244

```

Score = 76.3 bits (186), Expect = 1e-15
Identities = 42/92 (45%), Positives = 53/92 (56%), Gaps = 7/92 (7%)
Query: 148 YGCVPRQWCKFAAEITDPAK-----GNRVLIGTPTEDRALYDAAFMRGSA 200
          +*GVQPR WCK. AEE.P.A          RALWCTF TPA NALYDRAN. C
Sbjct: 70 PRGVQRWCKWAEITREPVRHGNSRSGKRLKLTPTATNAALYDRANSMVGYPTA 129

Query: 201 LNPPLRVNCEPDVPRIKSESSPSSNENG 232
          LNPP. G. +S
Sbjct: 130 RNFPPEDLGGRRKKDEASSGGVWLTNKG 161

```

>G1754 (Amino Acid Sequence) (gf=1)
Length = 341

Score = 75.9 bits (185), Expect = 2e-15
 identities = 49/126 (38%), Positives = 62/126 (48%), Gaps = 9/126 (7%)

Query: 127 PVKKETSPVSAVTAAKK-----HYRGVRPVGKPAEIDPANKARVWLGTF 178
 P K P + KKK YRGVROR WKG+ AEIR+P GAR+WLGT
 Sbjct: 43 PAKIRAPKSGKCGKCGKPGNGICDYGVRORRKGKVAIEP-DGCARLWLGT 101

Query: 187 AYDRAAFMRGSRALLNFP-----LRVNSGEPDPVRIKSKRS-----SPSSSN 229
A+D AA ++G A+LNFP R +S P +++ + ++ SPSSS+
Sbjct: 74 AHDVAALSIKASAILNFPDLAGSPRPSSLSRDIQVAALKAHMETSQSPSSSS 129

>G2134 (Amino Acid Sequence) (gf=1)
Length = 202
Score = 75.1 bits (183), Expect = 3e-15
Identities = 36/58 (62%), Positives = 42/58 (72%), Gaps = 1/58 (1%)

Query: 148 YRGVROPWPKFAEIRDPKNGARVWLGTPTETAEDALAYDRAAFMRGSRALLNFP 205
YRGVR+R WKG+ +EIR+P R+WLGTPTETAEDAA AYD AA AYD AAF RG A LNFP
Sbjct: 24 YRGVRRKRWKWSVSEIREPOTN-RIMLGSEFTEPMAATAYDAVAFHFRGREALNFP 80

>G977 (Amino Acid Sequence) (gf=1)
Length = 181
Score = 75.1 bits (183), Expect = 3e-15
Identities = 38/65 (58%), Positives = 44/65 (67%), Gaps = 1/65 (1%)

Query: 146 KHYGVROPWPKFAEIRDPKNGARVWLGTPTETAEDALAYDRAAFMRGSRALLNFP 205
+ +RGVROPWKG +EIR P R+WLGTPTETAEDAA AYD AA M G RA NFP
Sbjct: 6 QRGVROPWKGWSVSEIRHPLK-TRIMLGTFETAEDAAAYDEAARLMCGPARTNFP 64

Query: 206 LRVNS 210
N+
Sbjct: 65 YNPNA 69

>G24 (Amino Acid Sequence) (gf=1)
Length = 176
Score = 74.7 bits (182), Expect = 4e-15
Identities = 41/84 (48%), Positives = 52/84 (61%), Gaps = 10/84 (11%)

Query: 131 EKTSPVSAVTAAG-----KHVGVROPWPKFAEIRDPKNGARVWLGTPTETA 181
E + V+V A VTA K Y+G+R R WKG+ AEIR+P K +R+WLGT+ T
Sbjct: 2 ETEAVTATVTAATMGITGTRKEDLKPYGIRWKWKWVAEIREPNKR-SRINLGSTATP 60

Query: 182 EDALAYDRAAFMRGSRALLNFP 205
E AA AYD A F +RG A LNFP
Sbjct: 61 EAAARAYDTAVFYLRGSPARLNFP 84

>G10 (Amino Acid Sequence) (gf=1)
Length = 138
Score = 74.7 bits (182), Expect = 4e-15
Identities = 38/71 (53%), Positives = 48/71 (67%), Gaps = 5/71 (7%)

Query: 146 KHYGVROPWPKFAEIRDPKNGARVWLGTPTETAEDALAYDRAAFMRGSRALLNFP 205
K Y+G+R R WKG+ AEIR+P K +R+WLGT+TA AA AYD A F +RG A LNFP
Sbjct: 22 KRYGIRMRKWKWVAEIREPNKR-SRINLGSYKTAAVAAAYDAVAFYLRGSPARLNFP 80

Query: 206 LRV---NSGE 212
V N GE
Sbjct: 81 BEVPKDGNGGE 91

>G1387 (Amino Acid Sequence) (gf=1)
Length = 189
Score = 73.9 bits (180), Expect = 6e-15

Query: 179 ETEADALAYDRAAFMRGSRALLNFPRLVNSGEPDPVRIKSKRSPSSSNENGAKPKRR 238
++ +AALAYD AA + G A LN P N S + S+E+ +
Sbjct: 102 SSSYEALAYDEAAKAIYQASARLNLPETITNRSSSTAATATVSGSVTAFSDEVCARED 161

Query: 239 TVAAGG 244
T A+ G
Sbjct: 162 TNASSG 167

>G1090 (Amino Acid Sequence) (gf=1)
Length = 184
Score = 75.9 bits (185), Expect = 2e-15
Identities = 36/63 (57%), Positives = 46/63 (72%), Gaps = 1/63 (1%)

Query: 148 YRGVROPWPKFAEIRDPKNGARVWLGTPTETAEDALAYDRAAFMRGSRALLNFP 207
YRGVR+R WKG+ +EIR+P K R+WLGT+ETA E AA AYD AA +RG LNFP
Sbjct: 20 YRGIRERKWKWSVSEIREPKK-TRIMLGSEYETAEDAAAYDAALHLRGRTNLNFP 78

Query: 208 VNS 210
V+S
Sbjct: 79 VDS 81

>G1007 (Amino Acid Sequence) (gf=1)
Length = 225
Score = 75.9 bits (185), Expect = 2e-15
Identities = 43/83 (51%), Positives = 53/83 (63%), Gaps = 2/83 (2%)

Query: 123 VDSVVKKEKTSVSAVTAAGKHVGVROPWPKFAEIRDPKNGARVWLGTPTETA 182
VDS E +S T KG YRG R WKG+ +EIR+P K +R+WLGTPT TAE
Sbjct: 2 VDSHSGDTECSSKKKKEKTKGKGY-YRGARMSWKGWSVSEIREPKK-SRINLGTPPTAE 59

Query: 183 DAALAYDRAAFMRGSRALLNFP 205
AA A+D AA ++GS A+LNFP
Sbjct: 60 MAARAHDAVLAALSIKGSAILNFP 82

>G2299 (Amino Acid Sequence) (gf=1)
Length = 236
Score = 75.5 bits (184), Expect = 2e-15
Identities = 41/79 (51%), Positives = 54/79 (67%), Gaps = 4/79 (5%)

Query: 145 GKX--YRGVROPWPKFAEIRDPKNGARVWLGTPTETAEDALAYDRAAFMRGSRALL 202
GKH YRGVR R WKG+ +EIR+P K +R+WLGTPT T E AA A+D AA ++G+ A+L
Sbjct: 46 GKHPYRGVROPWPKFAEIRDPKNGARVWLGTPTETAEDALAYDRAAFMRGSRALL 104

Query: 203 NPELVNSGEPDPVRIKSK 221
NFP +S P PV + +
Sbjct: 105 NPELADS-FPRVLSLSPR 122

>G16 (Amino Acid Sequence) (gf=1)
Length = 218
Score = 75.5 bits (184), Expect = 2e-15
Identities = 48/116 (41%), Positives = 71/116 (60%), Gaps = 21/116 (18%)

Query: 129 KKEKTSVSAVTAAGKH--YRGVROPWPKFAEIRDPKNGARVWLGTPTETAEDAL 186
++EK PV + GKH YRGVR+R WKG+ +EIR+P K +R+WLGTPT + E AA
Sbjct: 20 EEEKKKPVKDS-----GKHPIYGVRRKRWKWSVSEIREPKK-SRINLGTFPSPENAR 73

Score = 72.0 bits (175), Expect = 2e-14
Identities = 34/58 (58%), Positives = 44/58 (75%), Gaps = 1/58 (1%)
Query: 148 YRGVROPWPKFAAEIRDPAKNGARVWLGTFTAEADAALAYDRAAFPMRGSRALLNFP 205
YRGVR R W GK+ +EIR+P K +R+WLGT+ TAE AA A+D AA ++G+ A LNFP
Sbjct: 100 YRGVRRMSGKWSEIREPRKK-SRIWLGTPTAEAAAHADVAALAKGTAYLNFP 156

>G2573 (Amino Acid Sequence) (gf=1)
Length = 307
Score = 71.6 bits (174), Expect = 3e-14
Identities = 38/74 (51%), Positives = 49/74 (65%), Gaps = 7/74 (9%)
Query: 130 KETSPVSAANTAAKHYGVROPWPKFAAEIRDPAKNGARVWLGTFTAEADAALAYD 189
K K P +A T +RGVROP W GK+ AEIR+P + G R+WLGT T + AA+AYD
Sbjct: 22 KGGGPNATCT-----PRGVORTWGWVAIEIRPNR-GTSLMGLTPTSVEAAMAYD 74

Query: 190 RAAPMRGSRALLN 203
AA ++ G A LN
Sbjct: 75 EAAKLYGHEAKLN 88

>G27 (Amino Acid Sequence) (gf=1)
Length = 179

Score = 71.6 bits (174), Expect = 3e-14
Identities = 39/79 (49%), Positives = 49/79 (61%), Gaps = 5/79 (6%)
Query: 131 EKTSPVSAANTAAK-----KHVROPWPKFAAEIRDPAKNGARVWLGTFTAEADAAL 186
EKT S KG +RGVR R W GK+ +EIR+P K +R+WLGT T E AA
Sbjct: 19 EKTGSKTKNEQKGYSPNPRGVRRMQWGWSEIREPRKK-SRIWLGTPTPEMAAR 77

Query: 187 AYDRAAFPMRGSRALLNFP 205
A+D AA ++G A LNFP
Sbjct: 78 AHDVAALAKGSAHLNFP 96

>G1381 (Amino Acid Sequence) (gf=1)
Length = 256

Score = 71.2 bits (173), Expect = 4e-14
Identities = 37/74 (50%), Positives = 49/74 (66%), Gaps = 2/74 (2%)
Query: 148 YRGVROPWPKFAAEIRDPAKNGARVWLGTFTAEADAALAYDRAAFPMRGSRALLNFP 207
YRGVR R W GK+ +EIR+P K +R+WLGT+ TAE AA A+D AA ++G+ LNFP
Sbjct: 71 YRGVRRMSGKWSEIREPRKK-SRIWLGTPTAEAAAHADVAALAKGSGFLNFP-E 128

Query: 208 VNSGEPDPVRIKX 221
++ P PV K
Sbjct: 129 LSGLLPRVSCSPK 142

>G2510 (Amino Acid Sequence) (gf=1)
Length = 192

Score = 70.9 bits (172), Expect = 5e-14
Identities = 40/95 (42%), Positives = 57/95 (59%), Gaps = 13/95 (13%)
Query: 123 VDSVPVKKEKTSVPVSAANTAAK-----KH--YRGVROPWPKFAAEIRDPAKNG 170
+ S PV + +AAV + G +H Y GVR+R W GK+ +EIR+P K
Sbjct: 8 LSSPVTNNEPTATASAVKSCGGGKETSSSTTRHPVYHGVKRWGWSEIREPRKK- 66

Query: 171 ARVILGTFTAEADALAYDRAAFPMRGSRALLNFP 205
+R+WLGT+P E AA AYD AAF ++G +A LNFP

Sbjct: 67 SRIWLQSPVPPEMAAKAYDVAAAFCLKGRKAQLNFP 101

>G2291 (Amino Acid Sequence) (gf=1)
Length = 256

Score = 70.5 bits (171), Expect = 7e-14
Identities = 34/57 (59%), Positives = 44/57 (76%), Gaps = 1/57 (1%)
Query: 148 YRGVROPWPKFAAEIRDPAKNGARVWLGTFTAEADAALAYDRAAFPMRGSRALLNFP 204
+ GVR+RPMG++AEIRD R WLGT+TAE+AA AYD AA R+RG++A NF
Sbjct: 116 PMGVRRPWRGWSAEIRDRIGR-CHRWLGTFTAEBAARAYDAARLRGTAKTNF 171

>G2108 (Amino Acid Sequence) (gf=1)
Length = 219

Score = 69.7 bits (169), Expect = 1e-13
Identities = 35/57 (61%), Positives = 42/57 (73%), Gaps = 1/57 (1%)
Query: 148 YRGVROPWPKFAAEIRDPAKNGARVWLGTFTAEADAALAYDRAAFPMRGSRALLNFP 204
Y GVR+RPMG++AEIR+P R WLGT+TAE+AA AYD AA + GS A NF
Sbjct: 21 YLGVRRPWRGVAEIRNPFYK-ERHWLGTFTAEBAAPAYDVAARSIGSLATTNF 76

>G3 (Amino Acid Sequence) (gf=1)
Length = 153

Score = 69.7 bits (169), Expect = 1e-13
Identities = 35/63 (55%), Positives = 44/63 (69%), Gaps = 1/63 (1%)
Query: 143 AKGHVYGVROPWPKFAAEIRDPAKNGARVWLGTFTAEADAALAYDRAAFPMRGSRALL 202
A K IRG+R R W GK+ AEIR+P K +R+WLGT+ T AA AYD A P +RG A L
Sbjct: 26 ATRPYRGIIRKRWKVAEIREPNKR-SRLWLGSYTTIDIAARAYDAVAYFIRGFSARL 84

Query: 203 NFP 205
NFP
Sbjct: 85 NFP 87

>G13 (Amino Acid Sequence) (gf=1)
Length = 253

Score = 68.9 bits (167), Expect = 2e-13
Identities = 38/77 (49%), Positives = 49/77 (63%), Gaps = 3/77 (3%)
Query: 130 KETSPVSAANTAAKGHYR--GVROPWPKFAAEIRDPAKNGARVWLGTFTAEADAALA 187
+ +T+P +KG + GVRORP GK+ AEI+D + R+WLGTPTAE+AA A
Sbjct: 2 EHQTTPKQTKKESKGNKTKFVGVRQPSGKWVAEIKDTTK-IRWLGTPTAEAAARA 60

Query: 188 YDRAAFPMRGSRALLNFP 204
YD AA +RGS NF
Sbjct: 61 YDEAACLLGNSNTRTNF 77

>G2115 (Amino Acid Sequence) (gf=1)
Length = 230

Score = 68.2 bits (165), Expect = 3e-13
Identities = 45/114 (39%), Positives = 57/114 (49%), Gaps = 25/114 (21%)
Query: 97 PSSSSDEDRSSPPVKIETPSPANDSVPVKKEKTSVPVSAANTAAKHYGVROPW 156
PSSSS SS S K K K S + K YGVR R W
Sbjct: 23 PSSSSSSSSSSSSSSCK-----NKNKSKI-----KRYKGYRRMSW 58

Gapped K H 0.140
Lambda 0.267 0.0410 0.140

Matrix: BLAOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 325807
Number of Sequences: 1566
Number of Extensions: 15185
Number of successful extensions: 306
Number of sequences better than 10.0: 151
Number of HSP's better than 10.0 without gapping: 148
Number of HSP's successfully gapped in prelim test: 3
Number of HSP's that attempted gapping in prelim test: 41
Number of HSP's gapped (non-prelim): 162
Length of query: 268
Length of database: 551,722
effective HSP length: 42
effective length of query: 226
effective length of database: 485,950
effective search space: 109824700
effective search space used: 109824700

T: 11
A: 40
X1: 16 (7.2 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 42 (21.9 bits)
S2: 49 (23.5 bits)

Query: 157 GKFAEIRDPKNGARVWLGTFTAEADALAYDRAAFMRGSRALLNPF 210
G + EIR P + R+WLG++ TAE AA AYD A ++G A LNPF +S

Sbjct: 59 GSWSEIRAPNQK-TRIMLGSYSTAEARAAYDVALLCKGQAMLNPTSSSS 111

>G869 (Amino Acid Sequence) (gf=1)
Length = 324

Score = 67.8 bits (164), Expect = 4e-13
Identities = 41/83 (49%), Positives = 53/83 (63%), Gaps = 5/83 (6%)

Query: 108 SPSPVKI-ETPESFANDVSPVKEKTSVSAATAKGHYGVGRORPWCKPAAEIRDP 166
+PFS+++ E P ++ DS K + VSA+ + K GYROR WCK+AAEIRDP

Sbjct: 75 NPFSEVSEPFSSSQSDST-KTDGKIIVASPAVPRKFPV-GYRQKNGKMAEIRDP 131

Query: 167 AKNGARVWLGTFTAEADALAYD 189
K R WLGTFT+T E+AA AYD

Sbjct: 132 IKK-TRTWLGTPTDLEAAKAYD 153

>G40 (Amino Acid Sequence) (gf=1)
Length = 213

Score = 67.4 bits (163), Expect = 6e-13
Identities = 34/57 (59%), Positives = 42/57 (73%), Gaps = 1/57 (1%)

Query: 148 YRGVQRPWCKPAAEIRDPKNGARVWLGTFTAEADALAYDRAAFMRGSRALLNPF 204
YRGVQR GK+ E+R+P K R+WLGTFT+TAE AA A+D AA +RG A LNPF

Sbjct: 48 YRGVQRNSGKWSVEENKK-TRIMLGTFTQTAEMARAHDVAALALRGSRACLNPF 103

>G41 (Amino Acid Sequence) (gf=1)
Length = 207

Score = 66.6 bits (161), Expect = 1e-12
Identities = 34/57 (59%), Positives = 41/57 (71%), Gaps = 1/57 (1%)

Query: 148 YRGVQRPWCKPAAEIRDPKNGARVWLGTFTAEADALAYDRAAFMRGSRALLNPF 204
YRGVQR GK+ E+R+P K R+WLGTFT+TAE AA A+D AA +RG A LNPF

Sbjct: 42 YRGVQRNSGKWCLEPNNKK-TRIMLGTFTQTAEMARAHDVAALALRGSRACLNPF 97

>G2067 (Amino Acid Sequence) (gf=1)
Length = 244

Score = 66.2 bits (160), Expect = 1e-12
Identities = 37/90 (41%), Positives = 48/90 (53%), Gaps = 1/90 (1%)

Query: 116 TPESFANDVSPVKEKTSVSAATAKGHYGVGRORPWCKPAAEIRDPKNGARVWL 175
T S + S TS + K K I+GVR R WG + EIR P + R+WL

Sbjct: 11 TSSSLSHSSSSSSSTLALRHOSCKKIKYKGVMRGSRVWTEIRAPNQK-TRIML 69

Query: 176 GTPEADAEALAYDRAAFMRGSRALLNPF 205
G++ TAE AA AYD A ++G A LNPF

Sbjct: 70 GSYSTAEARAAYDAALCLCKEKANLNPF 99

Database: ptfgene
Posted date: Aug 27, 2002 10:23 AM
Number of letters in database: 551,722
Number of sequences in database: 1566

Lambda	K	H
0.313	0.129	0.373